

Proceedings of the workshop

# HERA and the LHC

workshop series on the implications of HERA for LHC physics

2006 - 2008, Hamburg - Geneva

Part 1

Editors: Hannes Jung, Albert De Roeck

Verlag Deutsches Elektronen-Synchrotron

# Impressum

## Proceedings of the workshop HERA and the LHC

### 2nd workshop on the implications of HERA for LHC physics 2006 - 2008, Hamburg - Geneva

Conference homepage

<http://www.desy.de/~heralhc>

Online proceedings at

<http://www.desy.de/~heralhc/proceedings-2008/proceedings.html>

The copyright is governed by the Creative Commons agreement, which allows for free use and distribution of the articles for non-commercial activity, as long as the title, the authors' names and the place of the original are referenced.

Editors:

Hannes Jung (DESY, U. Antwerp), Albert De Roeck (CERN, U. Antwerp)

March 2009

DESY-PROC-2009-02

ISBN 978-3-935702-32-4

ISSN 1435-8077

Published by

Verlag Deutsches Elektronen-Synchrotron

Notkestraße 85

22607 Hamburg

Germany

## **Organizing Committee:**

H. Jung (DESY and U. Antwerp, chair), A. De Roeck (CERN and U. Antwerp, chair)  
G. Altarelli (CERN), J. Blümlein (DESY), M. Botje (Nikhef), J. Butterworth (UCL), K. Eggert (CERN),  
E. Gallo (INFN), T. Haas (DESY), R. Heuer (CERN), M. Klein (Liverpool), M. Mangano (CERN),  
A. Morsch (CERN), G. Polesello (INFN/CERN), O. Schneider (EPFL), C. Vallee (CPPM)

## **Conveners:**

### **Parton Density Functions:**

M. Dittmar (ETH, Zürich, CMS), S. Forte (U. Milan), A. Glazov (DESY, H1), S. Moch (DESY)

### **Multi-jet final states and energy flows:**

C. Gwenlan (UCL, ZEUS), L. Lönnblad (Lund), E. Rodrigues (LHCb), G. Zanderighi (CERN)

Contact persons: S. Banerjee (CMS), D. Traynor (H1)

### **Heavy quarks (charm and beauty):**

M. Cacciari (Paris VI & VII.), A. Dainese (INFN, ALICE), A. Geiser (DESY, ZEUS), H. Spiesberger (U. Mainz)

Contact persons: K. Lipka (U. Hamburg, H1), Ulrich Uwer (CERN)

### **Diffraction:**

M. Arneodo (U. Piemonte Orientale, Novara, INFN, CMS, ZEUS), M. Diehl (DESY), P. Newman (U. Birmingham, H1), V. A. Khoze (U. Durham)

Contact persons: A. Bruni (INFN, ZEUS), B. Cox (U. Manchester, ATLAS), R. Orava (U. Helsinki)

### **Cosmic Rays:**

C. Diaconu (DESY, CPPM, H1), Ch. Kiesling (MPI Munich, H1), T. Pierog (FZ Karlsruhe),

### **Monte Carlos and Tools:**

P. Bartalini (Taiwan, CMS), S. Chekanov (Argonne, ZEUS), F. Krauss (IPP Durham), S. Gieseke (U. Karlsruhe)

## **Advisory Committee:**

G. Altarelli (CERN), J. Bartels (Hamburg), M. Della Negra (CERN), J. Ellis (CERN), J. Engelen (CERN),  
G. Gustafson (Lund), G. Ingelman (Uppsala), P. Jenni (CERN), R. Klanner (DESY), M. Klein (DESY),  
L. McLerran (BNL), T. Nakada (CERN), D. Schlatter (CERN), F. Schrempp (DESY),  
J. Schukraft (CERN), J. Stirling (Durham), W.K. Tung (Michigan State), A. Wagner (DESY),  
R. Yoshida (ANL)

## **Supported by:**

Deutsches Elektronen-Synchrotron DESY (Hamburg)  
CERN (Geneva)

## Preface

In April 2004 the first meeting of what would become a series of HERA-LHC meetings took place at CERN. Over 250 participants joined and helped to shape the goals and objectives of this workshop. These are:

- To identify and prioritize those measurements to be made at HERA which have an impact on the physics reach of the LHC;
- to encourage and stimulate transfer of knowledge between the HERA and LHC communities and establish an ongoing interaction;
- to encourage and stimulate theory and phenomenology efforts;
- to examine and improve theoretical and experimental tools;
- to increase the quantitative understanding of the implication of HERA measurements on LHC physics;

That HERA deep inelastic scattering and photoproduction data and knowledge acquired will have an impact on the analysis of LHC data is a priori obvious. First and foremost there is the question on the structure of the proton. HERA is the first and so far only collider for lepton-proton scattering to date. The data from the 27.5 GeV electron beams scattered on the 820 (920) GeV protons have delivered an accurate picture of the Structure of the proton in a wide kinematic range. Precise predictions of cross sections at the LHC critically depend on the knowledge of the parton density functions (PDFs) in the proton. It can be the largest uncertainty in measurements, as the detector systematics will get under control to the anticipated level. The precision measurements at HERA in the last 15 years have boosted our knowledge on the parton distributions by several orders in magnitude in kinematic reach and by specific measurements of heavy flavors, such as bottom and charm quark PDFs. Final states allow to study multi-jet production, complementing the impressive LEP results by measurements in an environment with an additional important complication, namely in the presence of an object containing color: the proton. The understanding of these data will be a key to the study of LHC jet data especially at medium jet energies. Measurements of this type, together with the PDF data, allow for precision tests of QCD dynamics, e.g. to test classical approaches such as DGLAP evolution, or more sophisticated ideas, including e.g. special  $\log(1/x)$  terms, angular ordering etc. The HERA data are also instrumental in understanding double parton scattering, a phenomenon which is expected to be very important at the LHC. HERA has elevated the studies of diffraction to precision physics, and the LHC is expected to carry on that program. Finally, many tools have been developed over the last years for the analysis and understanding of HERA data, which can be adapted for future studies at the LHC.

In view of this anticipated synergy between HERA and the LHC the workshop has defined six working groups

- Parton density functions and related questions
- Multi-Jet final states and energy flows
- Heavy quarks (charm and beauty)
- Diffraction
- Cosmic Rays, HERA and the LHC
- Monte Carlo generators and tools

The *Parton Density Functions* working group had the most obvious task, namely getting to understand what the present precision - both from data and from theory- is to determine parton distributions, and what are the consequences of these uncertainties on LHC measurements. At an early stage in the workshop it became obvious that the combination of the H1 and ZEUS experiments would be very beneficial. Such lessons had been learned from LEP and are now applied at HERA. It turns out that the gain of a common analysis of the data of the two experiments on the precision of the PDFs is substantially larger than when these data-sets are used individually in fits. Benchmark test have been performed to check the systematics of the different assumptions in the QCD fit procedures, keeping certain assumptions and data sets in the fits fixed. At the start of the workshop there was some controversy on the NLO gluons at low- $x$ , being very different between different PDF fit groups. This could be resolved by measuring  $F_L$ , requiring lower energy running at HERA. The workshop has strongly supported that proposal and the last months of HERA have been used to measure  $F_L$ . First results are now being released by the experiments. Steps on getting towards common procedures to be used in the PDF fitting community and to get the most optimal PDFs are being defined and followed up in a special PDF forum called PDF4LHC, which is a spin-off of this workshop

The *Multi-Jet final states and energy flows* working group has studied in detail the novel jet algorithms, designed to be infrared and collinear safe, such as the SISCone and the (anti)- $k_T$  algorithms. Jet algorithms and performances as used in the experiments are discussed. A jet quality measure has been defined. The perturbative calculation of higher order corrections has been studied in detail and a comparison of all order analytical resummation with Monte Carlo parton shower approaches has been performed. An important issue to understand better the details of the final states in experimental data is the concept of  $k_T$  factorization. A formalism for extracting e.g. the needed unintegrated gluon distributions from fits to data is proposed. Implications for the LHC are studied e.g. on the case of gauge boson production, and boson production in association with heavy quarks. Forward so called "Mueller-Navalet" jets predictions have been made for the LHC. Very forward jet measurements opportunities e.g. using forward CASTOR detector in CMS look promising. Finally prompt photon production, high density systems and handles to multi-parton event discoveries have been discussed.

In the *Heavy Quarks (Charm and Beauty)* working group a summary of experimental results on fragmentation functions, gluon densities and charm/beauty masses from HERA has been collected. Prospects for heavy quark measurement at the LHC are discussed. In the theory area important and significant progress has been made in the understanding of heavy quark mass effects in the evolution of parton density functions. In a common contribution from members of CTEQ and MRST the progress in understanding of mass effects and its impact on the global analysis of parton density functions is reviewed and documented. In addition also progress in the calculation of fragmentation functions including mass effects is discussed. Finally the progress in calculation of higher order corrections to  $t\bar{t}$  production at the LHC is summarized.

The working group *Diffraction* brought about an important information transfer between HERA (and Tevatron) and the LHC on the experience with near beam detectors operation and calibration issues. Since the start of the workshop, there are several near beam detector projects that have been launched in the experiments. Diffractive and forward physics is now in the blood of the LHC experiments. CMS, TOTEM and ALICE present their physics program, also what can be achieved without near beam detectors by using rapidity gaps instead. Major progress has been achieved in understanding central exclusive production at  $pp$  colliders, with a tight re-evaluation of the theoretical calculations, and foremost with the exclusive measurements made at the Tevatron. Factorization in diffractive processes remains mysterious. It is known not to work between  $ep$  and  $pp$  data. Now also within  $ep$  data at HERA it is found not to work for diffractive di-jet photoproduction events. The deployment of diffractive PDFs from  $ep$  to  $pp$  data has therefore to be done with care.

In the working group *Cosmic Rays, HERA and the LHC* the impact of laboratory measurements for the understanding of the source and propagation of high energy cosmic rays has been discussed. These cosmic rays are measured mainly via air-showers and for their simulation measurements at high energy lepton hadron and hadron hadron colliders are important. The main sources of uncertainties come from cross sections (elastic and inelastic), secondary particle production and multiplicity distributions. Hadron production in the forward region especially from HERA and also the LHC can provide important constraints. On the theory side the application of perturbative QCD for the calculations, hadron production in the forward region, the relation to multi-parton interaction and non-linear effects arising at highest energies (i.e. at small

*x*) have been discussed.

The goals of working group *Monte Carlo and Tools* was to examine and improve the Monte Carlo event generators for the use at LHC, to provide a framework for tuning and to develop new tools and libraries for the analysis of data. The available Monte Carlo generators are reviewed and tools like HZtool and RIVET, tools for fitting like Professor and Proffit (in *Multi-Jet nal states and energy flows* WG) are discussed. Multiparton interaction and underlying event structures was a major issue, also in close connection with *Multi-Jet nal states and energy flows*.

The special character of this workshop was – apart from its clear charge on the connection between HERA and LHC – that it was alternative held at CERN and at DESY. Note that Tevatron was always an invited guest at the table, and its data and interpretation of the results have always been part of the input in the discussions.

The last workshop in this series was held at CERN, where the series started, and over 200 participants attended. This clearly shows that the workshop has been established to be a beacon and forum for discussions of QCD for the preparation of the LHC. With the termination of the HERA accelerator in 2007 and the turn on of the LHC, the series was terminated and the results are written up in these extensive proceedings. But clearly there is a need and community for targeted forum on LHC QCD questions, and no doubt a workshop of this kind will emerge in the near future, as soon as the first data arrive.

Finally we wish to thank all the participants of the HERA and the LHC workshops for making this series so interesting and lively. We thank especially the conveners for their enormous work in the preparation of the many meetings and finally the proceedings

Last but not least we wish to thank A. Grabowsky, D. Denise, S. Platz and L. Schmidt for their continuous help and support during all the meetings. We thank B. Liebaug for the design of the poster. We are grateful to R. Eisberg, O. Knak and S. König for recording the talks and all technical help. We thank M. Mayer, K. Sachs and M. Stein for their help in printing the proceedings. We are grateful to the CERN and DESY directorates for financial support of the workshops and for their encouragement to investigate the HERA - LHC connection in detail.

Hannes Jung and Albert De Roeck

# Contents

<b>1 WG: Parton Density Functions</b>	<b>1</b>
<b>Introduction</b>	<b>3</b>
M. Dittmar, S. Forte, A. Glazov, S. Moch, G. Altarelli, J. Anderson, R. D. Ball, G. Beuf, M. Boonekamp, H. Burkhardt, F. Caola, M. Ciafaloni, D. Colferai, A. Cooper-Sarkar, A. De Roeck, L. Del Debbio, J. Feltesse, F. Gelis, J. Grebenyuk, A. Guffanti, V. Halyo, J. I. Latorre, V. Lendermann, G. Li, L. Motyka, T. Petersen, A. Piccione, V. Radescu, M. Rogal, J. Rojo, C. Royon, G. P. Salam, D. Šálek, A. M. Stašo, R. S. Thorne, M. Ubiali, J. A. M. Vermaseren, A. Vogt, G. Watt, C. D. White	
<b>Theoretical issues</b>	<b>8</b>
S. Moch, M. Rogal, J. A. M. Vermaseren, A. Vogt, G. Altarelli, R. D. Ball, M. Ciafaloni, D. Colferai, G. P. Salam, A. Stašo, R. S. Thorne, C. D. White, G. Beuf, F. Caola, F. Gelis, L. Motyka, Ch. Royon, D. Šálek, A. M. Stašo	
<b>Benchmarking of parton distributions and their uncertainties</b>	<b>53</b>
R. D. Ball, L. Del Debbio, J. Feltesse, S. Forte, A. Glazov, A. Guffanti, J. I. Latorre, A. Piccione, V. Radescu, J. Rojo, R. S. Thorne, M. Ubiali, G. Watt	
<b>Determination of parton distributions</b>	<b>74</b>
A. Cooper-Sarkar, A. Glazov, G. Li, J. Grebenyuk, V. Lendermann	
<b>Proton-proton luminosity, standard candles and PDFs at the LHC</b>	<b>105</b>
J. Anderson, M. Boonekamp, H. Burkhardt, M. Dittmar, V. Halyo, T. Petersen	
<b>Outlook: the PDF4LHC initiative</b>	<b>125</b>
A. De Roeck	
<b>2 WG: Multi-Jet Final States and Energy Flows</b>	<b>127</b>
<b>Introduction</b>	<b>129</b>
C. Gwenlan, . Lönnblad, E. Rodrigues, G. Zanderighi (Eds.), A. Bacchetta, A. Banfi, S. Baranov, J. Bartels, A. Bunyatyan, V. Coco, G. Corcella, M. Dasgupta, M. Deák, P.-Antoine Delsart, I.M. Dremin, F. Hautmann, S. Joseph, H. Jung, A. Knutsson, K. Kutak, A. Lipatov, G. Luisoni, S. Majhi, L. Marti, K. Müller, T. Namsoo, S. Osman, H. Perrey, G. Rodrigo, J. Rojo, Z. Ruriková, A. Sabio Vera, C. Sander, Th. Schörner-Sadenius, F. Schwennsen, G. Somogyi, G.y Soye, M. Strikman, M. Treccani, D. Treleani, Z. Trócsányi, B.F.L. Ward, S.A. Yost, N. Zotov	
<b>Higher-order calculations</b>	<b>132</b>
G. Zanderighi, G. Rodrigo, M. Treccani, G. Somogyi	

<b>Event shapes and resummation</b>	<b>155</b>
A. Banfi, G. Corcella, M. Dasgupta, S. Joseph G. Luisoni, S. Majhi, B.F.L. Ward, S.A. Yost	
<b>Jets and jet algorithms</b>	<b>182</b>
V. Coco, P.-Antoine D., J. Rojo, Ch. Sander, G. Soyez	
<b><math>k_{\perp}</math>-factorization and forward jets</b>	<b>205</b>
S. Baranov, J. Bartels, M. Deák, F. Hautmann, H. Jung, A. Knutsson, K. Kutak, A. Lipatov, Ch. Royon, A. Sabio Vera, F. Schwennsen, N. Zotov	
<b>HERA Results</b>	<b>240</b>
K. Müller, H. Perrey, Th. Schörner-Sadenius	
<b>Interactions at high gluon densities</b>	<b>257</b>
M. Strikman, I. M. Dremin	
<b>3 WG: Heavy Quarks (Charm and Beauty)</b>	<b>271</b>
<b>Introduction</b>	<b>272</b>
M. Cacciari, A. Dainese, A. Geiser, H. Spiesberger	
<b>Experimental study of heavy flavour production at HERA</b>	<b>274</b>
S. Boutle, J. Bracinik, A. Geiser, G. Grindhammer, A.W. Jung, P. Roloff, Z. Rúriková, M. Turcato, A. Yagües-Molina	
<b>Experimental study of heavy flavour production at RHIC and LHC</b>	<b>296</b>
M. Biasini, C. Bombonati, G.E. Bruno, E. Lytken, A. Mischke, C. Rosemann, A. Starodumov, D. Stocco, R. Wolf, M. zur Nedden	
<b>Heavy flavour production at the LHC: Theoretical Aspects</b>	<b>331</b>
Z.J. Ajaltouni, A. Banfi, S. Baranov, I. Bierenbaum, J. Blümlein, G. Corcella, M. Czakon, G. Ferrera, S. Klein, B. A. Kniehl, G. Kramer, A. Likhoded, D. A. Milstead, O. I. Piskounova, V. A. Saleev, I. Schienbein, H. Spiesberger, R.S. Thorne, W.K. Tung, G. Zanderighi, N. Zotov	

# Chapter 1

## Working Group Parton Density Functions

### Convenors:

*M. Dittmar (ETH, Zürich, CMS),  
S. Forte (U. Milan),  
A. Glazov (DESY, H1)  
S. Moch (DESY)*



# Introduction

## CONVENERS:

*M. Dittmar*<sup>1</sup>, *S. Forte*<sup>2</sup>, *A. Glazov*<sup>3</sup>, *S. Moch*<sup>4</sup>

## CONTRIBUTING AUTHORS:

*G. Altarelli*<sup>5,6</sup>, *J. Anderson*<sup>7</sup>, *R. D. Ball*<sup>8</sup>, *G. Beuf*<sup>9</sup>, *M. Boonekamp*<sup>10</sup>, *H. Burkhardt*<sup>11</sup>,  
*F. Caola*<sup>2</sup>, *M. Ciafaloni*<sup>12</sup>, *D. Colferai*<sup>12</sup>, *A. Cooper-Sarkar*<sup>13</sup>, *A. de Roeck*<sup>14</sup>, *L. Del Debbio*<sup>8</sup>,  
*J. Feltesse*<sup>15,16</sup>, *F. Gelis*<sup>5</sup>, *J. Grebenyuk*<sup>3</sup>, *A. Guffanti*<sup>17</sup>, *V. Halyo*<sup>18</sup>, *J. I. Latorre*<sup>19</sup>,  
*V. Lendermann*<sup>20</sup>, *G. Li*<sup>21</sup>, *L. Motyka*<sup>22,23</sup>, *T. Petersen*<sup>14</sup>, *A. Piccione*<sup>2</sup>, *V. Radescu*<sup>3</sup>, *M. Rogal*<sup>4</sup>,  
*J. Rojo*<sup>2,24</sup>, *C. Royon*<sup>10</sup>, *G. P. Salam*<sup>24</sup>, *D. Šálek*<sup>25</sup>, *A. M. Staśto*<sup>26,27,28</sup>, *R. S. Thorne*<sup>29</sup>,  
*M. Ubiali*<sup>8</sup>, *J. A. M. Vermaseren*<sup>30</sup>, *A. Vogt*<sup>31</sup>, *G. Watt*<sup>29</sup>, *C. D. White*<sup>30</sup>

<sup>1</sup> Institute for Particle Physics, ETH-Zürich Höggerberg, CH 8093 Zürich, Switzerland

<sup>2</sup> Dipartimento di Fisica, Università di Milano, INFN Sezione di Milano, Via Celoria 16, I-20133 Milan, Italy

<sup>3</sup> DESY, Notkestrasse 85, D-22603 Hamburg, Germany

<sup>4</sup> DESY, Platanenallee 6, D-15738 Zeuthen, Germany

<sup>5</sup> CERN, Department of Physics, Theory Division, CH-1211 Geneva 23, Switzerland

<sup>6</sup> Dipartimento di Fisica “E. Amaldi”, Università Roma Tre and INFN, Sezione di Roma Tre, via della Vasca Navale 84, I-00146 Rome, Italy

<sup>7</sup> School of Physics, University College Dublin, Ireland

<sup>8</sup> School of Physics, University of Edinburgh, Edinburgh EH9 3JZ, UK

<sup>9</sup> Institut de Physique Théorique, CEA–Saclay F-91191 Gif-sur-Yvette, France

<sup>10</sup> IRFU, Service de Physique des Particules, CEA–Saclay F-91191 Gif-sur-Yvette, France

<sup>11</sup> CERN-AB, CH-1211 Geneva 23, Switzerland

<sup>12</sup> Dipartimento di Fisica, Università di Firenze and INFN, Sezione di Firenze, I-50019 Sesto Fiorentino, Italy

<sup>13</sup> Department of Physics, Nuclear and Astrophysics Lab., Keble Road, Oxford, OX1 3RH, UK

<sup>14</sup> CERN-PH, CH-1211 Geneva 23, Switzerland

<sup>15</sup> IRFU, CEA–Saclay F-91191 Gif-sur-Yvette, France

<sup>16</sup> University of Hamburg, Luruper Chaussee 149, Hamburg, D-22761 Germany

<sup>17</sup> Physikalisches Institut, Albert-Ludwigs-Universität Freiburg, Hermann-Herder-Straße 3, D-79104 Freiburg i. B., Germany

<sup>18</sup> Department of Physics, Princeton University, Princeton, NJ08544, USA

<sup>19</sup> Departament d’Estructura i Constituents de la Matèria, Universitat de Barcelona, Diagonal 647,

E 08028 Barcelona, Spain

<sup>20</sup> Kirchhoff-Institut für Physik, Universität Heidelberg, Im Neuenheimer Feld 227, D-69120 Heidelberg, Germany

<sup>21</sup> Laboratoire de l’Accélérateur Linéaire, Université Paris-Sud, Orsay, France

<sup>22</sup> II Institute for Theoretical Physics, University of Hamburg, Luruper Chaussee 149, Hamburg, D-22761 Germany

<sup>23</sup> Institute of Physics, Jagellonian University Reymonta 4, 30-059 Cracow, Poland

<sup>24</sup> LPTHE, UPMC – Paris 6, Paris-Diderot – Paris 7, CNRS UMR 7589, F-75005 Paris, France

<sup>25</sup> Institute of Particle and Nuclear Physics, Charles University, Prague, Czech Republic

<sup>26</sup> Physics Department, Penn State University, 104 Davey Laboratory, University Park, PA 16802, USA

<sup>27</sup> H. Niewodniczański Institute of Nuclear Physics, Polish Academy of Science, ul.Radzikowskiego 152, 31-342 Cracow, Poland

<sup>28</sup> Brookhaven National Laboratory, Upton, NY-11073, USA

<sup>29</sup> Department of Physics and Astronomy, University College, London, WC1E 6BT, UK

<sup>30</sup> NIKHEF Theory Group, Kruislaan 409, NL 1098 SJ Amsterdam, The Netherlands

<sup>31</sup> Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, UK

**Abstract**

We provide an assessment of the state of the art in various issues related to experimental measurements, phenomenological methods and theoretical results relevant for the determination of parton distribution functions (PDFs) and their uncertainties, with the specific aim of providing benchmarks of different existing approaches and results in view of their application to physics at the LHC.

We discuss higher order corrections, we review and compare different approaches to small  $x$  resummation, and we assess the possible relevance of parton saturation in the determination of PDFs at HERA and its possible study in LHC processes. We provide various benchmarks of PDF fits, with the specific aim of studying issues of error propagation, non-gaussian uncertainties, choice of functional forms of PDFs, and combination of data from different experiments and different processes. We study the impact of combined HERA (ZEUS-H1) structure function data, their impact on PDF uncertainties, and their implications for the computation of standard candle processes, and we review the recent  $F_L$  determination at HERA. Finally, we compare and assess methods for luminosity measurements at the LHC and the impact of PDFs on them.

With the start of data-taking at the LHC getting closer, the importance of a detailed understanding of the physics of parton distributions (PDFs) has increased considerably, along with the awareness of the LHC community for the importance of the issues related to it. Clearly, the main reason why PDFs are important at the LHC is that at a hadron collider a detailed understanding of PDFs is needed in order to obtain accurate predictions for both signal and background processes. Indeed, for many physical processes at the LHC, PDFs are the dominant source of uncertainty. On the other hand, an accurate control of PDF uncertainties allows one to use selected processes as “standard candles”, for instance in the determination of luminosities. However, this also means that experimentation at the LHC will provide a considerable amount of new experimental information on PDFs, and it will enable us to test the adequacy of their current theoretical understanding.

The main aim of this document is to provide a state of the art assessment of our understanding of PDFs at the dawn of the LHC. Since the previous HERA-LHC workshop [1], we have witnessed several important directions of progress in the physics of PDFs. On the theoretical side there has been conclusive progress in extending the treatment of perturbative QCD beyond the current default, namely, the next-to-leading perturbative order. On the phenomenological side there has been a joint effort between experimental and theoretical groups involved in the extraction of PDFs, specifically from global fits, in agreeing on common procedures, benchmarks and standards. On the experimental side, new improved results from the HERA runs are being finalized: these include both the construction of a joint determination of structure function which combines the result of the ZEUS and H1 experiments, and the first direct measurements of the structure function  $F_L$  which have been made possible by running HERA at a reduced proton beam energy in 2007. Also, the LHC experiments (ATLAS, CMS and LHCb) are now assessing

the use of standard candle processes for luminosity measurements.

All these issues are discussed in this document. In each case, our main goal has been to provide as much as possible a joint treatment by the various groups involved, as well as a comparison of different approaches and benchmarking of results. In particular, in Sect. [2], after briefly reviewing (Sect. [3]) the current status of higher-order calculations for DIS, we provide (Sect. [4]) detailed comparisons of techniques and results of different existing approaches to small  $x$  resummation, and then we summarize (Sect. [5]) the current status of studies of parton saturation at HERA, their possible impact on current PDF extraction and the prospects of future studies at the LHC. In Sect. [6] we discuss methods and results for the benchmarking of PDF fits: with specific reference to two benchmark fits based on a common agreed set of data, we discuss issues related to error propagation and non-gaussian errors, to the choice of functional form and corresponding bias, to possible incompatibilities between different data sets. In Sect. [7] we turn to recent progress in the extraction of PDFs from HERA data, specifically the impact of combined ZEUS-H1 structure function data on PDF determination and the ensuing calculation of  $W$  and  $Z$  cross-sections (Sect. [8]) and the recent first determination of the structure function  $F_L$  (Sect. [9]). In Sect. [10] we discuss and compare luminosity measurements based on absolute proton-proton luminosity measurements to those based on the use of standard candle processes, and the impact on all of them of PDF uncertainties. Finally, in Sect. [11] we present the PDF4LHC initiative, which will provide a framework for the continuation of PDF studies for the LHC.

*Note:* Most of the contributions to this workshop are the result of collaboration between various groups. The common set of authors given for each section or subsection has read and approved the entire content of that section or subsection; however, when a subset of these authors is given for a specific part of the section or subsection, they are responsible for it.

## ACKNOWLEDGEMENTS

This work was supported in part by the following grants and agencies: the European network HEPTOOLS under contract MRTN-CT-2006-035505; ANR-05-JCJC-0046-01 (France) PRIN-2006 (Italy); MEC FIS2004-05639-C02-01; (Spain) and the Scottish Universities Physics Alliance (UK).

## References

- [1] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.
- [2] S. Moch *et al.*, *Theoretical issues*. These proceedings.
- [3] S. Moch *et al.*, *Precision calculations for inclusive dis: an update*. These proceedings.
- [4] G. Altarelli *et al.*, *Small  $x$  resummation*. These proceedings.
- [5] G. Beuf *et al.*, *Parton saturation and geometric scaling*. These proceedings.
- [6] R. D. Ball *et al.*, *Benchmarking of parton distributions and their uncertainties*. These proceedings.

- [7] A. Cooper-Sarkar et al, *Determination of parton distributions*. These proceedings.
- [8] A. Cooper-Sarkar, A. Glazov, G. Li, *Extraction of the proton pdfs from a combined fit of  $hI$  and zeus inclusive dis cross sections*. These proceedings.
- [9] J. Grebenyuk, V. Lendermann, *Measurements of the proton structure function  $f_1$  at hera*. These proceedings.
- [10] J. Anderson et al , *Proton–proton luminosity, standard candles and pdfs at the lhc*. These proceedings.
- [11] A. de Roeck, *Outlook: the pdf4lhc initiative*. These proceedings.

# Theoretical issues

*S. Moch, M. Rogal, J. A. M. Vermaseren, A. Vogt, G. Altarelli, R. D. Ball, M. Ciafaloni, D. Colferai, S. Forte, G. P. Salam, A. Stařto, R. S. Thorne, C. D. White, G. Beuf, F. Caola, F. Gelis, L. Motyka, C. Royon, D. řálek, A. M. Stařto*

## 1 Precision calculations for inclusive DIS: an update<sup>1</sup>

With high-precision data from HERA and in view of the outstanding importance of hard scattering cross sections at the LHC, a quantitative understanding of deep-inelastic processes is indispensable, necessitating calculations beyond the standard next-to-leading order of perturbative QCD.

In this contribution we briefly discuss the recent extension of the three-loop calculations for inclusive deep-inelastic scattering (DIS) [1–8] to the complete set of coefficient functions for the charged-current (CC) case. The new third-order expressions are too lengthy for this short overview. They can be found in Refs. [9, 10] together with the calculational methods and a more detailed discussion. Furthermore the reader is referred to Refs. [11, 12] for our first results on the three-loop splitting functions for the evolution of helicity-dependent parton distributions.

Structure functions in inclusive deep-inelastic scattering are among the most extensively measured observables. The combined data from fixed-target experiments and the HERA collider spans about four orders of magnitude in both Bjorken- $x$  variable and the scale  $Q^2 = -q^2$  given by the momentum  $q$  of the exchanged electroweak gauge boson [13]. Here we consider the  $W$ -exchange charged-current case, see Refs. [14–20] for recent data from neutrino DIS and HERA. With six structure functions,  $F_2^{W^\pm}$ ,  $F_3^{W^\pm}$  and  $F_L^{W^\pm}$ , this case has a far richer structure than, for example, electromagnetic DIS with only two independent observables,  $F_2$  and  $F_L$ .

Even taking into account a forthcoming combined H1/ZEUS final high- $Q^2$  data set from HERA, more detailed measurements are required to fully exploit the resulting potential, for instance at a future neutrino factory, see Ref. [21], and the LHeC, the proposed high-luminosity electron-proton collider at the LHC [22]. Already now, however, CC DIS provides important information on the parton structure of the proton, e.g., its flavour decomposition and the valence-quark distributions. Moreover, present results are also sensitive to electroweak parameters of the Standard Model such as  $\sin^2 \theta_W$ , see Ref. [23], and the space-like  $W$ -boson propagator [24]. As discussed, for example, in Refs. [25–28], a reliable determination of  $\sin^2 \theta_W$  from neutrino DIS requires a detailed understanding of non-perturbative and perturbative QCD effects.

Previous complete results on unpolarized DIS include the three-loop splitting functions [4, 5] as well as the 3-loop coefficient functions for the photon-exchange structure functions  $F_{2,L}$  [6, 7]. However, most coefficient functions for CC DIS were not fully computed to three loops so far.

For this case it is convenient to consider linear combinations of the structure functions  $F_a^{W^\pm}$  with simple properties under crossing, such as  $F_a^{\nu p \pm \bar{\nu} p}$  ( $a = 2, 3, L$ ) for neutrino DIS.

---

<sup>1</sup>Contributing authors: S. Moch, M. Rogal, J. A. M. Vermaseren, A. Vogt

For all these combinations either the even or odd moments can be calculated in Mellin- $N$  space in the framework of the operator product expansion (OPE), see Ref. [29]. The results for the third-order coefficient functions for the even- $N$  combinations  $F_{2,L}^{\nu p+\bar{\nu}p}$  can be taken over from electromagnetic DIS [6, 7]. Also the coefficient function for the odd- $N$  based charged-current structure function  $F_3^{\nu p+\bar{\nu}p}$  is completely known at three-loop accuracy, with the results only published via compact parameterizations so far [8]. For the remaining combinations  $F_{2,L}^{\nu p-\bar{\nu}p}$  and  $F_3^{\nu p-\bar{\nu}p}$ , on the other hand, only recently the first six odd or even integer moments of the respective coefficient functions have been calculated to third order in Ref. [9] following the approach of Refs. [1–3] based on the MINCER program [30, 31].

The complete results of Refs. [6–8] fix all even and odd moments  $N$ . Hence already the present knowledge of fixed Mellin moments for  $F_{2,L}^{\nu p-\bar{\nu}p}$  and  $F_3^{\nu p-\bar{\nu}p}$  is sufficient to determine also the lowest six moments of the differences of corresponding even- $N$  and odd- $N$  coefficient functions and to address a theoretical conjecture [32] for these quantities, see Ref. [10]. Furthermore these moments facilitate  $x$ -space approximations in the style of, e.g. Ref. [33] which are sufficient for most phenomenological purposes, including the determination of the third-order QCD corrections to the Paschos-Wolfenstein relation [34] used for the extraction of  $\sin^2 \theta_W$  from neutrino DIS.

The even-odd differences of the CC coefficient functions  $C_a$  for  $a = 2, 3, L$  can be defined by

$$\delta C_{2,L} = C_{2,L}^{\nu p+\bar{\nu}p} - C_{2,L}^{\nu p-\bar{\nu}p}, \quad \delta C_3 = C_3^{\nu p-\bar{\nu}p} - C_3^{\nu p+\bar{\nu}p}. \quad (1)$$

The signs are chosen such that the differences are always ‘even – odd’ in the moments  $N$  accessible by the OPE [29], and it is understood that the  $d^{abc}d_{abc}$  part of  $C_3^{\nu p+\bar{\nu}p}$  [3, 8] is removed before the difference is formed. With  $a_s = \alpha_s/(4\pi)$  these non-singlet quantities can be expanded as

$$\delta C_a = \sum_{l=2} a_s^l \delta c_a^{(l)}. \quad (2)$$

There are no first-order contributions to these differences, hence the above sums start at  $l = 2$ .

We start the illustration of these recent results by looking at the approximations for the  $\nu p - \bar{\nu}p$  odd- $N$  coefficient functions  $c_{2,L}^{(3)}(x)$  (see Ref. [10] for a detailed discussion). These are compared in Fig. 1 to their exact counterparts [6, 7] for the even- $N$  non-singlet structure functions. The dashed lines represent the uncertainty band due to the limited number of known moments. The third-order even-odd differences remain noticeable to larger values of  $x$  than at two loops, e.g., up to  $x \simeq 0.3$  for  $F_2$  and  $x \simeq 0.6$  for  $F_L$  for the four-flavour case shown in the figure. The moments  $N = 1, 3, \dots, 9$  constrain  $\delta c_{2,L}^{(3)}(x)$  very well at  $x \gtrsim 0.1$ , and approximately down to  $x \approx 10^{-2}$ .

Concerning low values of Bjorken- $x$  one should recall that the uncertainty bands shown by the dashed lines in Fig. 1 do not directly indicate the range of applicability of these approximations, since the coefficient functions enter observables only via smoothing Mellin convolutions with non-perturbative initial distributions. In Fig. 2 we therefore present the convolutions of all six third-order CC coefficient functions with a characteristic reference distribution. It turns

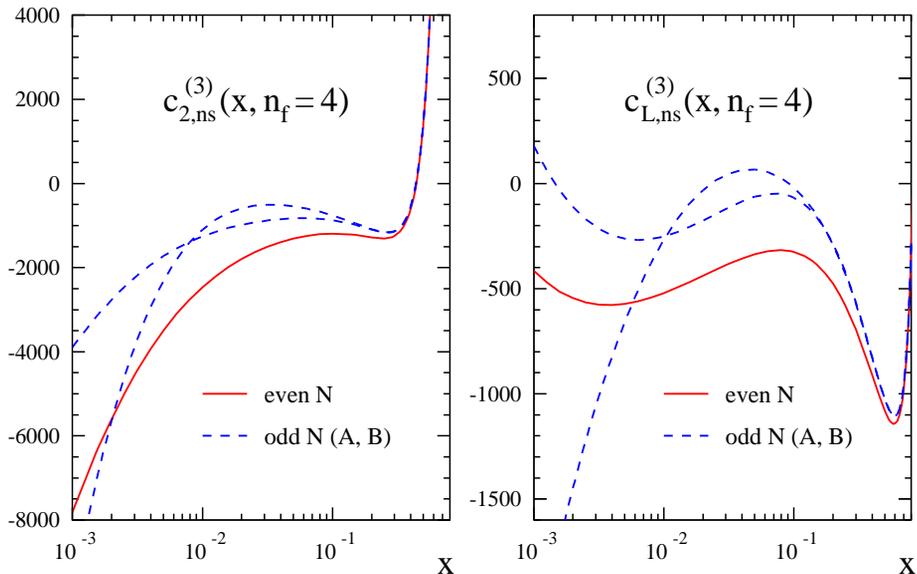


Fig. 1: The exact third-order coefficient functions of the even- $N$  structure functions  $F_{2,L}^{\nu p + \bar{\nu} p}$  for four massless flavours, and the approximate odd-moment quantities for  $\nu p - \bar{\nu} p$  combination.

out that the approximations of the previous figure can be sufficient down to values even below  $x = 10^{-3}$ , which is amply sufficient for foreseeable applications to data. The uncertainty of  $\delta c_3^{(3)}(x)$ , on the other hand, becomes relevant already at larger values,  $x \lesssim 10^{-2}$ , as the lowest calculated moment of this quantity,  $N = 2$ , has far less sensitivity to the behaviour at low  $x$ .

The three-loop corrections to the non-singlet structure functions are rather small even well below the  $x$ -values shown in the figure – recall our small expansion parameter  $\alpha_s$ : the third-order coefficient are smaller by a factor  $2.0 \cdot 10^{-3}$  if the expansion is written in powers of  $\alpha_s$ . Their sharp rise for  $x \rightarrow 1$  is understood in terms of soft-gluon effects which can be effectively resummed, if required, to next-to-next-to-next-to-leading logarithmic accuracy [35]. Our even-odd differences  $\delta c_a^{(3)}(x)$ , on the other hand, are irrelevant at  $x > 0.1$  but have a sizeable impact at smaller  $x$  in particular on the corrections for  $F_2$  and  $F_L$ . The approximate results for  $\delta c_a^{(3)}(x)$  facilitate a first assessment of the perturbative stability of the even-odd differences (1). In Fig. 3 we illustrate the known two orders for  $F_2$  and  $F_L$  for  $\alpha_s = 0.25$  and  $n_f = 4$  massless quark flavours, employing the same reference quark distribution as in Fig. 2.

Obviously our new  $\alpha_s^3$  corrections are important wherever these coefficient-function differences are non-negligible. On the other hand, our results confirm that these quantities are very small, and thus relevant only when a high accuracy is required. These conditions are fulfilled for the calculation of QCD corrections for the so-called Paschos-Wolfenstein relation. This relation is defined in terms of a ratio of neutral-current and charged-current cross sections for neutrino-nucleon DIS [34],

$$R^- = \frac{\sigma(\nu_\mu N \rightarrow \nu_\mu X) - \sigma(\bar{\nu}_\mu N \rightarrow \bar{\nu}_\mu X)}{\sigma(\nu_\mu N \rightarrow \mu^- X) - \sigma(\bar{\nu}_\mu N \rightarrow \mu^+ X)}. \quad (3)$$

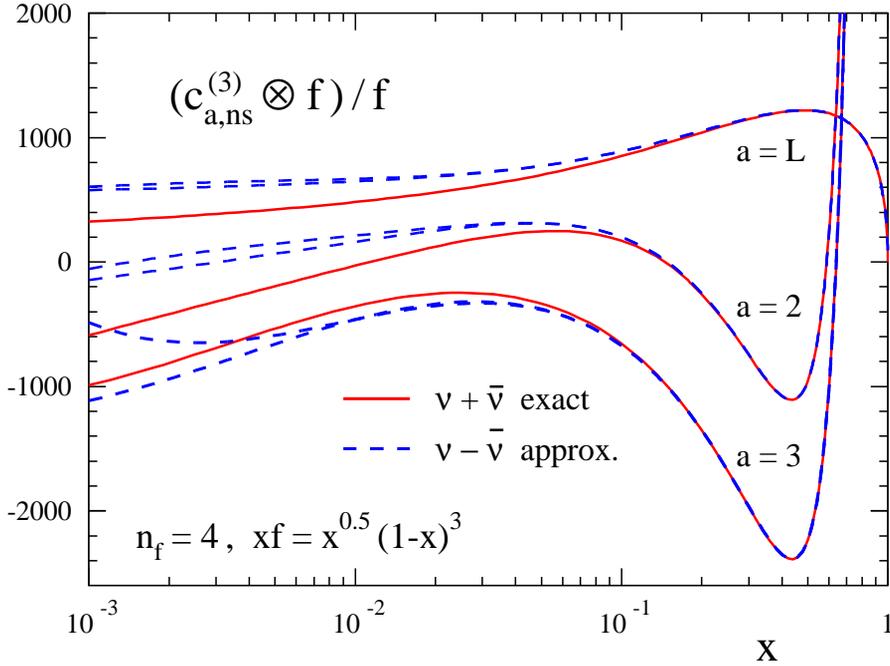


Fig. 2: Convolution of the six third-order CC coefficient functions for  $F_{2,3,L}$  in  $\nu p + \bar{\nu} p$  and  $\nu p - \bar{\nu} p$  DIS with a schematic but typical non-singlet distribution  $f$ . All results have been normalized to  $f(x)$ , suppressing the large but trivial variation of the absolute convolutions.

The asymmetry  $R^-$  directly measures  $\sin^2 \theta_W$  if the up and down valence quarks in the target carry equal momenta, and if the strange and heavy-quark sea distributions are charge symmetric. Beyond the leading order this asymmetry can be presented as an expansion in  $\alpha_s$  and inverse powers of the dominant isoscalar combination  $u^- + d^-$ , where  $q^- = \int_0^1 dx x (q(x) - \bar{q}(x))$  is the second Mellin moment of the valence quark distributions. Using the results for differences  $\delta c_a^{(3)}(x)$ ,  $a = 2, L, 3$  one can present it in a numeric form,

$$R^- = \frac{1}{2} - \sin^2 \theta_W + \frac{u^- - d^- + c^- - s^-}{u^- + d^-} \left\{ 1 - \frac{7}{3} \sin^2 \theta_W + \left( \frac{1}{2} - \sin^2 \theta_W \right) \cdot \frac{8 \alpha_s}{9 \pi} [1 + 1.689 \alpha_s + (3.661 \pm 0.002) \alpha_s^2] \right\} + \mathcal{O}((u^- + d^-)^{-2}) + \mathcal{O}(\alpha_s^4), \quad (4)$$

where the third term in the square brackets is determined by the  $\alpha_s^3$  corrections  $\delta c_a^{(3)}(x)$ ,  $a = 2, L, 3$ . The perturbation series in the square brackets appears reasonably well convergent for relevant values of the strong coupling constant, with the known terms reading, e.g.,  $1 + 0.42 + 0.23$  for  $\alpha_s = 0.25$ . Thus the  $\alpha_s^2$  and  $\alpha_s^3$  contributions correct the NLO estimate by 65% in this case. On the other hand, due to the small prefactor of this expansion, the new third-order term increases the complete curly bracket in Eq. (4) by only about 1%, which can therefore be considered as the new uncertainty of this quantity due to the truncation of the perturbative expansion. Consequently previous NLO estimates of the effect of, for instance, the (presumably

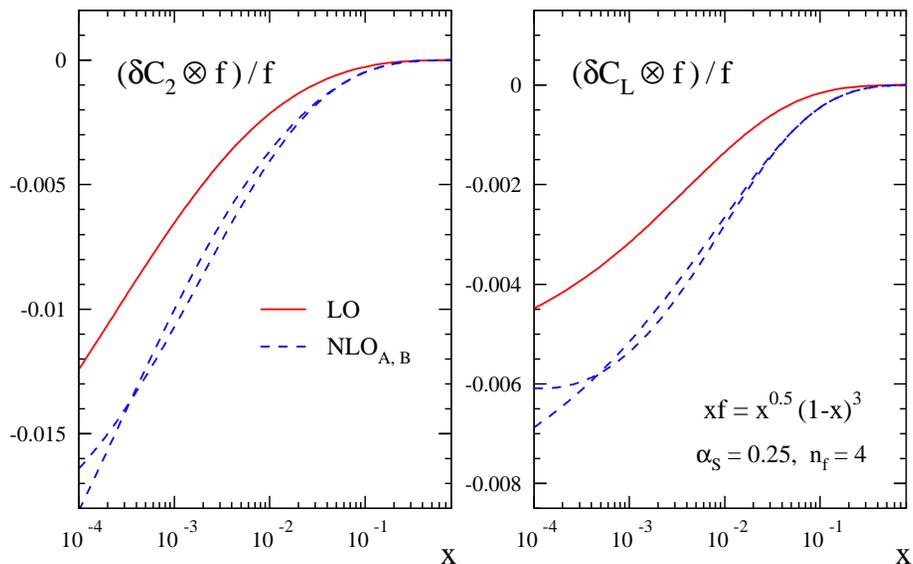


Fig. 3: The first two approximations, denoted by LO and NLO, of the differences (1) for  $F_2$  and  $F_L$  in charged-current DIS. The results are shown for representative values of  $\alpha_s$  and  $n_f$  after convolution with the reference distribution  $f(x)$  also employed in Fig. 2. The dashed curves correspond to the two approximation uncertainties for the new  $\alpha_s^3$  contributions.

mainly non-perturbative, see Refs. [36–38]) charge asymmetry of the strange sea remain practically unaffected by higher-order corrections to the coefficient functions.

To summarize, we have extended the fixed- $N$  three-loop calculations of inclusive DIS [1–3] to all charged-current cases not covered by the full (all- $N$ ) computations of Refs. [6–8]. The region of applicability of these new results is restricted to Bjorken- $x$  values above about  $10^{-3}$ , a range amply sufficiently for any fixed-target or collider measurements of those charged-current structure functions in the foreseeable future. Except for the longitudinal structure function  $F_L$ , the present coefficient functions are part of the next-to-next-to-next-to-leading order ( $N^3$ LO) approximation of massless perturbative QCD. Analyses at this order are possible outside the small- $x$  region since the corresponding four-loop splitting functions will have a very small impact here, cf. Ref. [39].

## 2 Small $x$ resummation <sup>2</sup>

The splitting functions which govern the evolution of the parton distributions (PDFs), together with the hard cross sections which relate those partons to hadronic physical observables, are potentially unstable at high energy due to logarithmically enhanced contributions. In particular, parametrizing observables such as deep-inelastic structure (DIS) functions or Drell-Yan (DY) or Higgs production cross section in hadronic collisions in terms of a dimensionful scale  $Q^2$  (photon virtuality or invariant mass of the final state in DIS and DY respectively) and a dimensionless ratio

<sup>2</sup>Contributing authors: G. Altarelli, R. D. Ball, M. Ciafaloni, D. Colferai, S. Forte, G. P. Salam, A. Stařto, R. S. Thorne, C. D. White

$x$  (the Bjorken variable or  $\frac{Q^2}{s}$  in DIS and DY respectively), when  $x \rightarrow 0$  there are logarithmically enhanced contributions to the perturbation expansion of the form  $x^{-1}\alpha_S^n(Q^2)\log^m(1/x)$  ( $n \geq m - 1$ ). When  $x$  is sufficiently small, one must resum such terms, reordering the perturbation expansion in terms of leading logarithmic (LL) terms followed by next-to-leading logarithmic (NLL) terms and so on.

The problem can be traced to ladders of  $t$ -channel gluon exchanges at LL order, with some quark mixing at NLL order and beyond. The underlying framework for the resummation procedure is the BFKL equation [40, 41], an integral equation for the unintegrated gluon  $f(k^2, Q_0^2)$  that is currently known up to full NLL order [42–44], and approximate NNLL order [45]. This has the schematic form (up to NLL):

$$Nf(k^2, Q_0^2) = Nf_I(Q_0^2) + \bar{\alpha}_S(k^2) \int dk'^2 \left[ \mathcal{K}_0(k^2, k'^2, Q_0^2) + \bar{\alpha}_S(k^2)\mathcal{K}_1(k^2, k'^2, Q_0^2) \right] f(k'^2), \quad (5)$$

where  $f_I(Q_0^2)$  is a non-perturbative initial condition at some initial scale  $Q_0$ ,  $\bar{\alpha}_S = 3\alpha_S/\pi$  and  $\mathcal{K}_{0,1}$  are the LL and NLL BFKL kernels. Different choices for the argument of the running coupling are possible, leading to accordingly modified  $\mathcal{K}_1$  [46, 47].

The solution of the BFKL equation can be used to extract leading and subleading singular contributions to singlet DGLAP splitting functions. The BFKL equation can either be solved numerically in its form given by Eq. (5), or else analytically by performing a double Mellin transform with respect to  $x$  and  $k^2$ :

$$f(\gamma, N) = \int_0^\infty (k^2)^{-\gamma-1} \int_0^1 dx x^N f(x, k^2), \quad (6)$$

whereby the BFKL equation becomes a differential equation, with kernels  $\chi_{0,1}(\gamma)$  defined respectively as the Mellin transforms of  $\mathcal{K}_{0,1}$ . Furthermore, by using the  $k_t$ -factorisation theorem [48], one may determine leading small  $x$  contributions to all orders to hard partonic cross sections for physical processes such as heavy quark electroproduction [48] and deep-inelastic scattering [49]. Approximate subleading results are also available [50, 51].

These results for splitting functions and hard partonic cross sections can then be combined with fixed-order results to obtain resummed predictions for physical observables. However, it has now been known for some time that the LL BFKL equation is unable to describe scattering data well, even when matched to a fixed order expansion. Any viable resummation procedure must then, at the very least, satisfy the following requirements:

1. Include a stable solution to the BFKL equation with running coupling up to NLL order.
2. Match to the standard DGLAP description at moderate and high  $x$  values (where this is known to describe data well).
3. Provide the complete set of splitting and coefficient functions for  $F_2$  and  $F_L$  in a well defined factorisation scheme.

Over the past few years, three approaches have emerged which, to some extent, aim at fulfilling these conditions. Here we call these the ABF [52–59], CCSS [47, 60–66] and TW [67–72] approaches. In the ABF scheme all three requirements are met, and resummed splitting functions in the singlet sector have been determined. Furthermore, a complete control of the

scheme dependence at the resummed level has been achieved, thereby allowing for a consistent determination of resummed deep-inelastic coefficient functions, and thus of resummed structure functions. However, the results obtained thus have not been fit to the data yet. In the CCSS formalism, resummed splitting functions have also been determined. However, results are given in a scheme which differs from the  $\overline{\text{MS}}$  scheme at the resummed level; furthermore, resummed coefficient functions and physical observables haven't been constructed yet. The TW approach, instead, has already been compared to the data in a global fit. However, this approach makes a number of simplifying assumptions and the ensuing resummation is thus not as complete as that which obtains in other approaches: for example, this approach does not include the full collinear resummation of the BFKL kernel.

A comparison of resummed splitting functions and solution of evolution equations determined in the ABF and CCSS approaches with  $n_f = 0$  was presented in Ref. [73]; the main features and differences of these approaches were also discussed. Here, we extend this comparison to the case of  $n_f \neq 0$  resummation, and also to the TW approach. First, we will briefly summarize the main features of each approach, and in particular we display the matrix of splitting functions determined in the ABF and CCSS approaches. Then, we will compare  $K$ -factors for physical observables determined using the ABF and TW approach.

Note that there are some difference in notations between various groups, which are retained here in order to simplify comparison to the original literature. In particular, the variable  $N$  in Eq. (6) will be referred to as  $\omega$  in the CCS approach of Section 2.2, and the variable  $\gamma$  in the same equation will be referred to as  $M$  in the ABF approach of Section 2.1.

## 2.1 The Altarelli-Ball-Forte (ABF) Approach

In the ABF approach [52–59, 74–77] one concentrates on the problem of obtaining an improved anomalous dimension (splitting function) for DIS which reduces to the ordinary perturbative result at large  $N$  (large  $x$ ), thereby automatically satisfying renormalization group constraints, while including resummed BFKL corrections at small  $N$  (small  $x$ ), determined through the renormalization-group improved (i.e. running coupling) version of the BFKL kernel. The ordinary perturbative result for the singlet anomalous dimension is given by:

$$\gamma(N, \alpha_s) = \alpha_s \gamma_0(N) + \alpha_s^2 \gamma_1(N) + \alpha_s^3 \gamma_2(N) \dots \quad (7)$$

The BFKL corrections at small  $N$  (small  $x$ ) are determined by the BFKL kernel  $\chi(M, \alpha_s)$ :

$$\chi(M, \alpha_s) = \alpha_s \chi_0(M) + \alpha_s^2 \chi_1(M) + \dots, \quad (8)$$

which is the Mellin transform, with respect to  $t = \ln \frac{k^2}{k_0^2}$ , of the  $N \rightarrow 0$  angular averaged BFKL kernel.

The ABF construction is based on three ingredients.

1. *The duality relation* between the kernels  $\chi$  and  $\gamma$

$$\chi(\gamma(N, \alpha_s), \alpha_s) = N, \quad (9)$$

which is a consequence of the fact that at fixed coupling the solutions of the BFKL and DGLAP equations should coincide at leading twist [52, 74, 78]. By using duality, one

can use the perturbative expansions of  $\gamma$  and  $\chi$  in powers of  $\alpha_s$  to improve (resum) each other: by combining them, one obtains a "double leading" (DL) expansion which includes all leading (and subleading, at NLO) logs of  $x$  and  $Q^2$ . In particular, the DL expansion automatically resums the collinear poles of  $\chi$  at  $M = 0$ . This eliminates the alternating sign poles  $+1/M, -1/M^2, \dots$  that appear in  $\chi_0, \chi_1, \dots$ , and make the perturbative expansion of  $\chi$  unreliable. This result is a model independent consequence of momentum conservation  $\gamma(1, \alpha_s) = 0$ , whence, by duality:

$$\chi(0, \alpha_s) = 1. \quad (10)$$

2. *The symmetry of the BFKL kernel* upon gluon interchange. In Mellin space, this symmetry implies that at the fixed-coupling level the kernel  $\chi$  for evolution in  $\ln \frac{s}{k k_0}$  must satisfy  $\chi(M) = \chi(1 - M)$ . By exploiting this symmetry, one can use the collinear resummation of the region  $M \sim 0$  which was obtained using the double-leading expansion to also improve the BFKL kernel in the anti-collinear  $M \simeq 1$  region. This leads to a symmetric kernel which is an entire function for all  $M$ , and has a minimum at  $M = \frac{1}{2}$ . The symmetry is broken by the DIS choice of variables  $\ln \frac{1}{x} = \ln \frac{s}{Q^2}$  and by the running of the coupling; however these symmetry breaking contribution can be determined exactly. This then leads to a stable resummed expansion of the resummed anomalous dimension at the fixed coupling level.
3. *The running-coupling resummation* of the BFKL solution. Whereas running coupling corrections to evolution equations are automatically included when solving the DGLAP evolution equation with resummed anomalous dimensions, the duality relation Eq. (9) itself undergoes corrections when the running coupling is included in the BFKL equation (5). Running coupling corrections can then be derived order by order, and turn out to be affected by singularities in Mellin  $M$  space. This implies that after Mellin inversion the associate splitting functions is enhanced as  $x \rightarrow 0$ : their contribution grows as  $(\alpha_s \beta_0 \ln \frac{1}{x})^n$  with the perturbative order. However the series of leading enhanced contribution can be summed at all orders in closed form, because it corresponds to the asymptotic expansion in powers of  $\alpha_s$  of the solution to the running coupling BFKL equation (5) when the kernel  $\chi$  is approximated quadratically about its minimum. This exact solution can be expressed in terms of Airy functions [53, 79] when the kernel is linear in  $\alpha_s$  and in terms of Bateman [55] functions for generic kernels. Because both the exact solution and its asymptotic expansion are known, this BFKL running coupling resummation can be combined with the DGLAP anomalous dimension, already resummed at the BFKL fixed coupling level, with full control of overlap (double counting terms). Schematically, the result has the following form:

$$\begin{aligned} \gamma_{\Sigma NLO}^{rc}(\alpha_s(t), N) &= \gamma_{\Sigma NLO}^{rc, pert}(\alpha_s(t), N) + \gamma^B(\alpha_s(t), N) - \gamma_s^B(\alpha_s(t), N) \\ &\quad - \gamma_{ss}^B(\alpha_s(t), N) - \gamma_{ss,0}^B(\alpha_s(t), N) + \gamma_{\text{match}}(\alpha_s(t), N) + \gamma_{\text{mom}}(\alpha_s(t), N), \end{aligned} \quad (11)$$

where  $\gamma_{\Sigma NLO}^{rc, pert}(\alpha_s(t), N)$  contains all terms which are up to NLO in the double-leading expansion of point 1, symmetrized as discussed in point 2 above so that its dual  $\chi$  has a minimum;  $\gamma^B(\alpha_s(t), N)$  resums the series of singular running coupling corrections using the aforementioned exact BFKL solution in terms of a Bateman function;  $\gamma_s^B(\alpha_s(t), N)$ ,

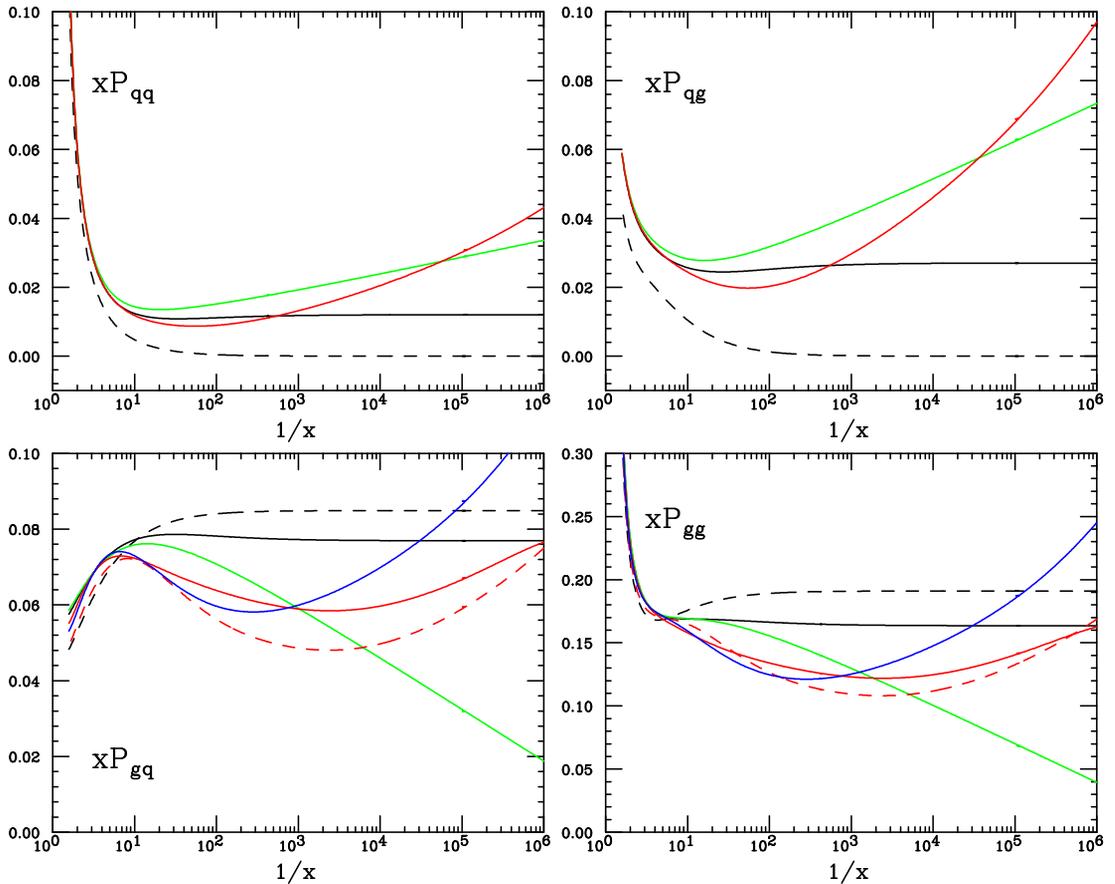


Fig. 4: The resummed splittings functions  $P_{qq}$ ,  $P_{qg}$ ,  $P_{gq}$  and  $P_{gg}$  in the ABF approach, all for  $n_f = 4$  and  $\alpha_s = 0.2$ : LO DGLAP (dashed black), NLO DGLAP (solid black), NNLO DGLAP (solid green), LO resummed (red dashed), NLO resummed in the  $\text{Q}_0\overline{\text{MS}}$  scheme (red) and in the  $\overline{\text{MS}}$  scheme (blue).

$\gamma_{ss}^B(\alpha_s(t), N)$   $\gamma_{ss,0}^B(\alpha_s(t), N)$  are double counting subtractions between the previous two contributions;  $\gamma_{\text{mom}}$  subtracts subleading terms which spoil exact momentum conservation;  $\gamma_{\text{match}}$  subtracts any contribution which deviates from NLO DGLAP and at large  $N$  doesn't drop at least as  $\frac{1}{N}$ .

The anomalous dimension obtained through this procedure has a simple pole as a leading small- $N$  (i.e. small  $x$ ) singularity, like the LO DGLAP anomalous dimension. The location of the pole is to the right of the DGLAP pole, and it depends on the value of  $\alpha_s$ . Thanks to the softening due to running of the coupling, this value is however rather smaller than that which corresponds to the leading BFKL singularity: for example, for  $\alpha_s = 0.2$ , when  $n_f = 0$  the pole is at  $N = 0.17$ .

The splitting function obtained by Mellin inversion of the anomalous dimension eq. (11)

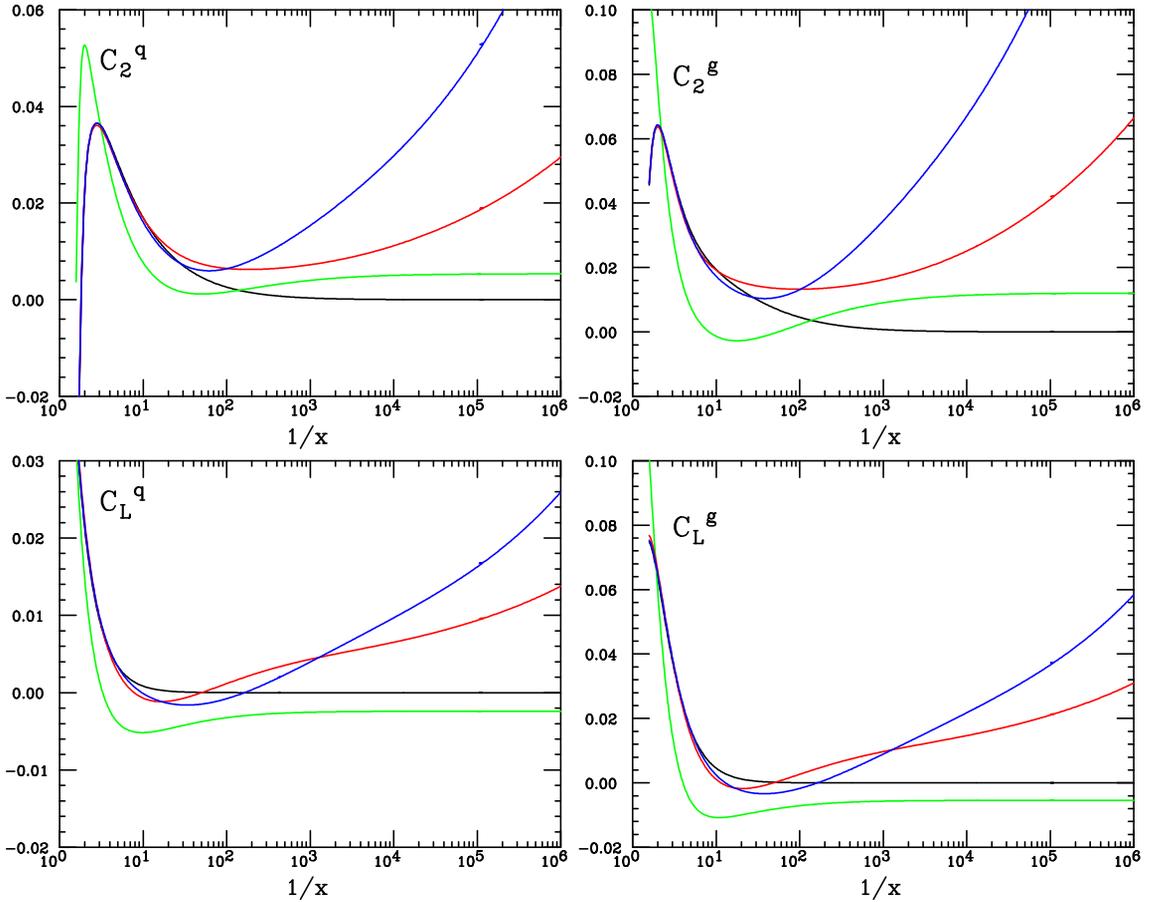


Fig. 5: The resummed DIS coefficient functions  $C_{2q}$ ,  $C_{2g}$ ,  $C_{Lq}$  and  $C_{Lg}$  in the ABF approach, all for  $n_f = 4$  and  $\alpha_s = 0.2$ . The curves are labelled as in the previous figure.

turns out to agree at the percent level to that obtained by the CCSS group by numerical resolution of the BFKL equation for all  $x \lesssim 10^{-2}$ ; for larger values of  $x$  (i.e. in the matching region) the ABF result is closer to the NLO DGLAP result.

In order to obtain a full resummation of physical observables, specifically for deep-inelastic scattering, the resummation discussed so far has to be extended to the quark sector and to hard partonic coefficients. This, on top of various technical complications, requires two main conceptual steps:

- A factorization scheme must be defined at a resummed level. Because only one of the two eigenvectors of the matrix of anomalous dimensions is affected by resummation, once a scheme is chosen, the resummation discussed above determines entirely the two-by-two matrix of splitting functions in the singlet sector. The only important requirement is that the relation of this small  $x$  scheme choice to standard large  $x$  schemes be known exactly,

since this enables one to combine resummed results with known fixed order results.

- PDFs evolved using resummed evolution equations must be combined with resummed coefficient functions. These are known, specifically for DIS [49], but are also known [80] to be affected by singularities, analogous to the running coupling singularities of the resummed anomalous dimension discussed above, which likewise must be resummed to all orders [57]. This running coupling resummation of the coefficient function significantly softens the small  $x$  growth of the coefficient function and substantially reduces its scheme dependence [58].

These steps have been accomplished in Ref. [58], where resummed anomalous dimensions (see fig. 4), coefficient functions (see fig.5) and structure functions (see section 2.4 below) have been determined. The scheme dependence of these results can be studied in detail: results have been produced and compared in both the  $\overline{\text{MS}}$  and  $\text{Q}_0\overline{\text{MS}}$  schemes, and furthermore the variation of results upon variation of factorization and renormalization scales has been studied.

Calculations of resummation corrections not only of deep inelastic processes, but also of benchmark hadronic processes such as Drell-Yan, vector boson, heavy quark and Higgs production are now possible and should be explored.

## 2.2 The Ciafaloni-Colferai-Salam-Stasto (CCSS) Approach

The Ciafaloni-Colferai-Salam-Stasto (CCSS) resummation approach proposed in a series a papers [47, 60–66] is based on the few general principles:

- We impose the so-called kinematical constraint [81–83] onto the real gluon emission terms in the BFKL kernel. The effect of this constraint is to cut out the regions of the phase space for which  $k_T'^2 \geq k_T^2/z$  where  $k_T, k_T'$  are the transverse momenta of the exchanged gluons and  $z$  is the fraction of the longitudinal momentum.
- The matching with the DGLAP anomalous dimension is done up to the next-to-leading order.
- We impose the momentum sum rule onto the resummed anomalous dimensions.
- Running coupling is included with the appropriate choice of scale. We take the argument of the running coupling to be the transverse momentum squared of the emitted gluon in the BFKL ladder in the BFKL part. For the part which multiplies the DGLAP terms in the eigenvalue equation we choose the scale to be the maximal between  $k_T^2$  and  $k_T'^2$ .
- All the calculations are performed directly in momentum space. This in particular enables easy implementation of the running of the coupling with the choice of the arguments as described above.

The implementation at the leading logarithmic level in BFKL and DGLAP (and in the single gluon channel case) works as follows. It is convenient to go to the Mellin space representation where we denote by  $\gamma$  and  $\omega$  the Mellin variables conjugated to  $\ln k_T$  and  $\ln 1/x$  respectively. The full evolution kernel can be represented as a series  $\mathcal{K} = \sum_n \alpha_s^{n+1} \mathcal{K}_n(\gamma, \omega)$ . We take the resummed kernel at the lowest order level to be

$$\mathcal{K}_0(\gamma, \omega) = \frac{2C_A}{\omega} \chi_0^\omega(\gamma) + [\gamma_0^{gg}(\omega) - \frac{2C_A}{\omega}] \chi_c^\omega(\gamma). \quad (12)$$

The terms in (12) are the following

$$\chi_0^\omega(\gamma) = 2\psi(1) - \psi(\gamma) - \psi(1 - \gamma + \omega) ,$$

is the leading logarithmic BFKL kernel eigenvalue with the kinematical constraint imposed. This is reflected by the fact that the singularities in the  $\gamma$  plane at  $\gamma = 1$  are shifted by the  $\omega$ . This ensures the compatibility with the DGLAP collinear poles, in the sense that we have only single poles in  $\gamma$ . The function  $\chi_c(\gamma)$  is the collinear part of the kernel

$$\chi_c^\omega(\gamma) = \frac{1}{\gamma} + \frac{1}{1 - \gamma + \omega} ,$$

which includes only the leading collinear poles at  $\gamma = 0$  or  $1$ . All the higher twist poles are neglected for this part of the kernel. This kernel eigenvalue is multiplied by the non-singular (in  $\omega$ ) part of the DGLAP anomalous dimension  $\gamma_0^{gg}(\omega) - 2C_A/\omega$  where  $\gamma_0^{gg}(\omega)$  is the full anomalous dimension at the leading order. The next-to-leading parts both in BFKL and DGLAP are included in the second term in the expansion, i.e. kernel  $\mathcal{K}_1$

$$\mathcal{K}_1(\gamma, \omega) = \frac{(2C_A)^2}{\omega} \tilde{\chi}_1^\omega(\gamma) + \tilde{\gamma}_1^{gg}(\omega) \chi_c^\omega(\omega) \quad (13)$$

where  $\tilde{\chi}_1^\omega(\gamma)$  is the NLL in x part of the BFKL kernel eigenvalue with subtractions. These subtractions are necessary to avoid double counting: we need to subtract the double and triple collinear poles in  $\gamma$  which are already included in the resummed expression (12) and which can be easily identified by expanding this expression in powers of  $\omega$  and using the LO relation  $\omega = \bar{\alpha}_s \chi_0(\gamma)$ . The term  $\tilde{\gamma}_1^{gg}(\omega)$  in Eq. (13) is chosen so that one obtains the correct DGLAP anomalous dimension at a fixed next-to-leading logarithmic level. The formalism described above has been proven to work successfully in the single channel case, that is for evolution of gluons only. The solution was shown to be very stable with respect to the changes of the resummation scheme.

The quarks are included in the CCSS approach by a matrix formalism. The basic assumptions in this construction are:

- Consistency with the collinear matrix factorization of the PDFs in the singlet evolution.
- Requirement that only single pole singularities in both in  $\gamma$  and  $\omega$  are present in the kernel eigenvalues. This assumption allows for the natural consistency with DGLAP and BFKL respectively. Higher order singularities can be generated at higher orders only through the subleading dependencies on these two variables.
- Ability to compute all the anomalous dimensions which can be directly compared with the DGLAP approach. This can be done by using set of recursive equations which allow to calculate the anomalous dimensions order by order from the kernel eigenvalues.
- Impose the collinear-anticollinear symmetry of the kernel matrix via the similarity transformation.
- Incorporate NLLx BFKL and DGLAP up to NLO (and possibly NNLO).

The direct solutions to the matrix equations are the quark and gluon Green's functions. These are presented in Fig. 6 for the case of the gluon-gluon and quark-gluon part. The resulting

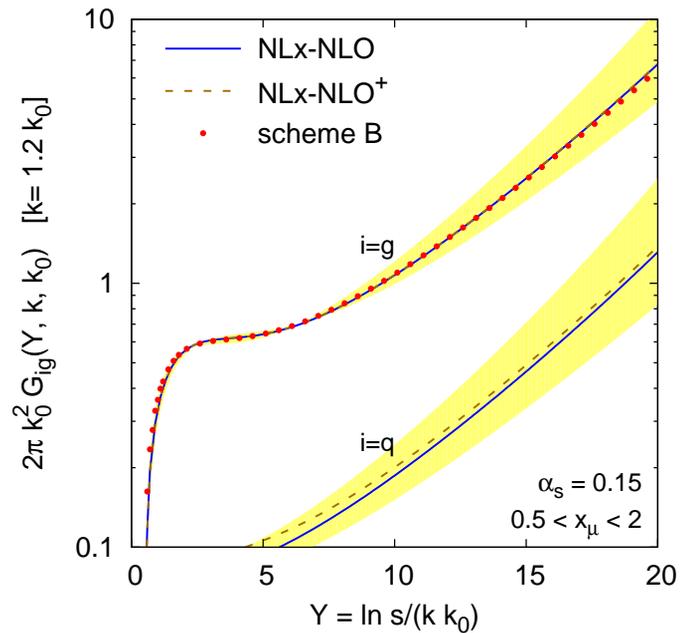


Fig. 6: Gluon-induced part of the Green function for the  $NLx\text{-NLO}$  and  $NLx\text{-NLO}^+$  models, compared to the results the single channel approach. For the models of this paper both gluon-gluon and quark-gluon Green's function are shown. The value chosen for the coupling,  $\alpha_s = 0.15$ , corresponds to  $k_0 \simeq 20 \text{ GeV}$ . The band indicates the spread in the result for the  $NLx\text{-NLO}$  model when varying the renormalization scale in the range  $0.5 < x_\mu < 2$ .

gluon-gluon part is increasing exponentially with the logarithm of energy  $\ln s$  with an effective intercept of about  $\sim 0.25$ . It is much suppressed with respect to the leading logarithmic order. We also note that the single channel results and the matrix results for the gluon-gluon Green's function are very similar to each other. In Fig. 6 we also present the quark-gluon channel which is naturally suppressed in normalization with respect to the gluon-gluon one by a factor of the strong coupling constant. This can be intuitively understood as the (singlet) quarks are radiatively generated from the gluons, and therefore this component follows the gluon density very closely. The yellow bands indicate the change of the Green's functions with respect to the change of the scale.

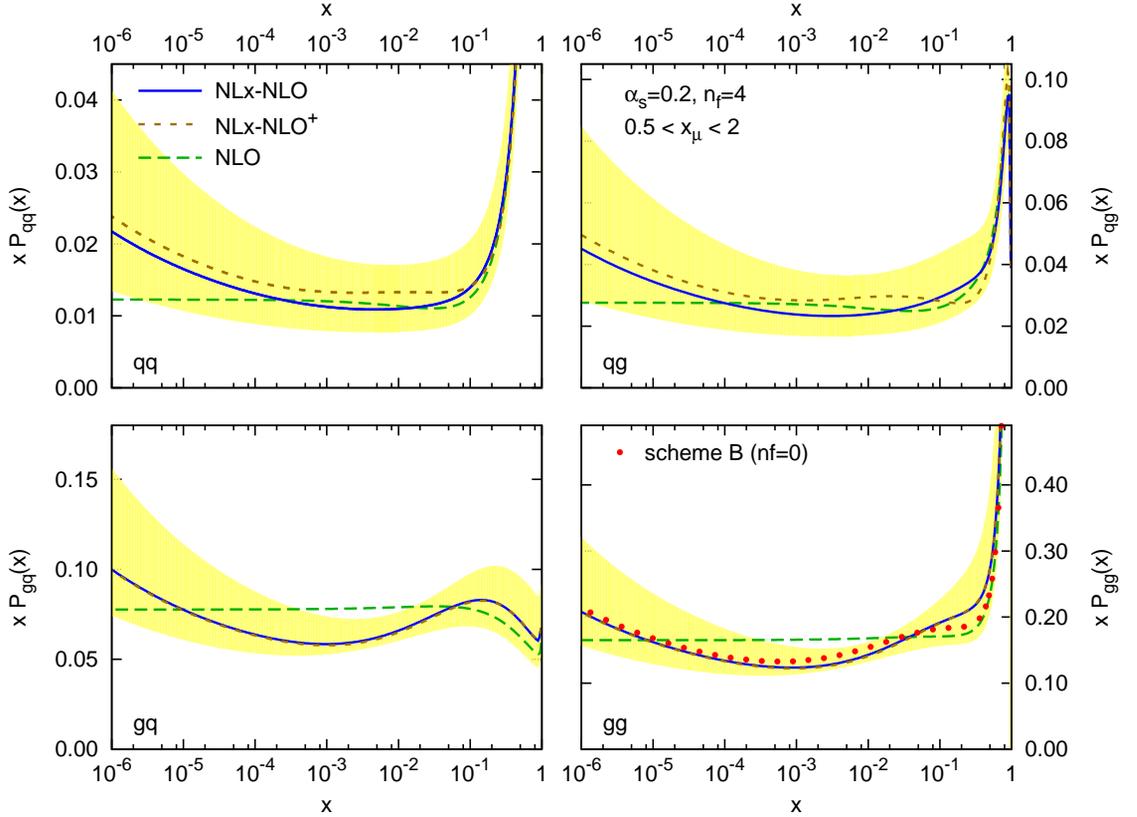


Fig. 7: The matrix of  $NLx-NLO$  (and  $NLx-NLO^+$ ) splitting functions together with their scale uncertainty and the  $NLO$  splitting functions for comparison. In the  $gg$  channel, we also show the old scheme B result ( $n_f = 0$ , no  $NLO$  contributions, 1-loop coupling). The band corresponds to the span of results ( $NLx-NLO$ ) obtained if one chooses  $x_\mu = 0.5$  and  $x_\mu = 2.0$ .

In Fig. 7 we present all four splitting functions for fixed value of scale  $Q^2$ . Here, again the results are very close to the previous single channel approach in the case of the gluon-gluon splitting function. The gluon-quark channel is very close to the gluon-gluon one, with the characteristic dip of this function at about  $x \sim 10^{-3}$ . The dip delays the onset of rise of the splitting function only to values of  $x$  of about  $10^{-4}$ . The scale dependence grows with decreasing  $x$  but it is not larger than in the fixed  $NLO$  case. The quark-gluon and quark-quark splitting functions

tend to have slightly larger uncertainty due to the scale change but are also slightly closer to the plain NLO calculation. They also tend to have a less pronounced dip structure.

### 2.3 The Thorne-White (TW) Approach

Substituting the LO running coupling  $\bar{\alpha}_S(k^2)$  into equation (5) and performing a double Mellin transform according to equation (6), the BFKL equation 5, as mentioned in Section 2, becomes a differential equation:

$$\frac{d^2 f(\gamma, N)}{d\gamma^2} = \frac{d^2 f_I(\gamma, Q_0^2)}{d\gamma^2} - \frac{1}{\beta_0 N} \frac{d(\chi_0(\gamma) f(\gamma, N))}{d\gamma} + \frac{\pi}{3\beta_0^2 N} \chi_1(\gamma) f(\gamma, N), \quad (14)$$

where  $\chi_{0,1}(\gamma)$  are the Mellin transforms of  $\mathcal{K}_{0,1}$ . The solution for  $f(N, \gamma)$  of Eq. (14) has the following form [61, 84]:

$$f(N, \gamma) = \exp\left(-\frac{X_1(\gamma)}{\beta_0 N}\right) \int_{\gamma}^{\infty} A(\tilde{\gamma}) \exp\left(\frac{X_1(\tilde{\gamma})}{\beta_0 N}\right) d\tilde{\gamma}. \quad (15)$$

Up to power-suppressed corrections, one may shift the lower limit of the integral  $\gamma \rightarrow 0$ , so that the gluon distribution factorises into the product of a perturbative and a non-perturbative piece. The nonperturbative piece depends on the bare input gluon distribution and an in principle calculable hard contribution. However, this latter part is rendered ambiguous by diffusion into the infrared, and in this approach is contaminated by infrared renormalon-type contributions. The perturbative piece is safe from this and is sensitive to diffusion into the ultraviolet region of weaker coupling. Substituting equation (15) into (14), one finds that the perturbative piece is given (after transforming back to momentum space):

$$\mathcal{G}_E^1(N, t) = \frac{1}{2\pi i} \int_{1/2-i\infty}^{1/2+i\infty} \frac{f^{\beta_0}}{\gamma} \exp[\gamma t - X_1(\gamma, N)/(\beta_0 N)] d\gamma, \quad (16)$$

where:

$$X_1(\gamma, N) = \int_{\frac{1}{2}}^{\gamma} \left[ \chi_0(\tilde{\gamma}) + N \frac{\chi_1(\tilde{\gamma})}{\chi_0(\tilde{\gamma})} \right] d\tilde{\gamma}. \quad (17)$$

Structure functions  $F_i$  also factorize, and the perturbative factors have a similar form to Eq. (16), but involve an additional impact factor  $h_i(\gamma, N)$  in the integrand according to the  $k_t$ -factorisation theorem [49]. Crucially, coefficient functions and anomalous dimensions involve ratios of the above quantities, such that the non-perturbative factor cancels. Thus, once all the impact factors are known, the complete set of coefficient and splitting functions can be disentangled. Finally they can be combined with the standard NLO DGLAP results (which are known to describe data well at higher  $x$  values) using the simple prescription:

$$P^{tot.} = P^{NLL} + P^{NLO} - \left[ P^{NLL(0)} + P^{NLL(1)} \right], \quad (18)$$

where  $P$  is a splitting or coefficient function, and  $P^{NLL(i)}$  the  $\mathcal{O}(\alpha_s^i)$  contribution to the resummed result which is subtracted to avoid double-counting. It should be noted that the method

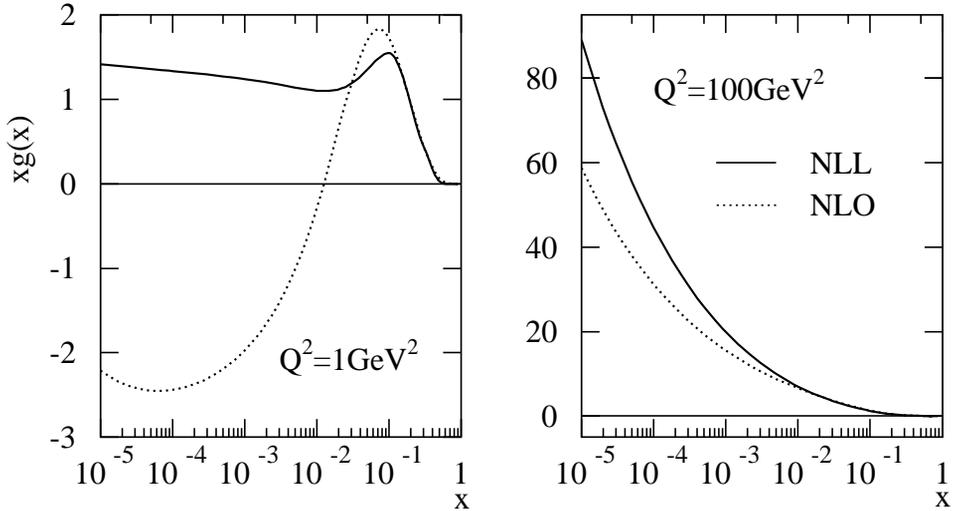


Fig. 8: Gluons arising from a global fit to scattering data including NLL small  $x$  resummations in the  $\text{DIS}(\chi)$  factorisation scheme (solid). Also shown is the result from an NLO DGLAP fit in the same scheme.

of subtraction of the resummed contribution in the matching is different to that for the ABF approach outlined after Eq. (11). For example, at NLO in the resummation the BFKL equation provides both the  $\alpha_S/N$  part of  $P_{gg}$  and the part at  $\mathcal{O}(\alpha_S)$  constant as  $N \rightarrow \infty$ . Hence we choose to keep all terms constant as  $N \rightarrow \infty$  generated by Eq. (16), with similar considerations for other splitting functions and coefficient functions, though these can contain terms  $\propto N$ . Hence, we include terms which will have some influence out to much higher  $x$  than in the ABF approach.

In the TW manner of counting orders LL is defined as the first order at which contributions appear, so while for the gluon splitting function this is for  $\bar{\alpha}_S^n \ln^m(1/x)$  for  $m = n - 1$  for impact factors this is for  $m = n - 2$ . A potential problem therefore arises in that the NLL impact factors are not known exactly. However, the LL impact factors with conservation of energy of the gluon imposed are known in cases of both massless and massive quarks [50, 51], and are known to provide a very good approximation to the full  $\mathcal{O}(\alpha_S^2)$  and  $\mathcal{O}(\alpha_S^3)$  quark-gluon splitting functions and coefficient functions [85], implying that they must contain much of the important higher-order information. These can then be used to calculate NLL coefficient and splitting functions within a particular factorisation scheme. One must also specify a general mass variable number scheme for consistent implementation of heavy quark mass effects. Such a scheme (called the  $\text{DIS}(\chi)$  scheme) has been given in [71, 72] up to NLL order in the high energy expansion, and

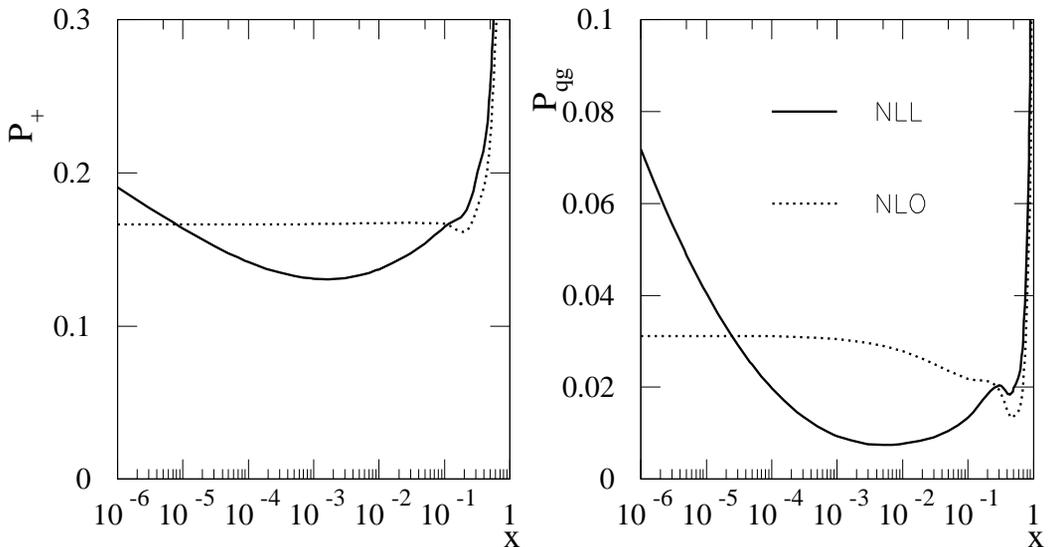


Fig. 9: The resummed splitting functions (solid)  $P_+ \approx P_{gg}$  and  $P_{qg}$  in the TW approach, both for  $n_f = 4$  and  $\alpha_s = 0.16$ , compared to the corresponding NLO forms (dotted).

NLO order in the fixed order expansion.

The form of the resummed splitting functions shown in fig. 9 are qualitatively consistent with those from the ABF approach, fig. 4, and CCSS approach fig. 7 (note however that in these plots the value of  $\alpha_s$  is a little larger, and the scheme is different). This is despite the fact that the approach does not include the explicit collinear resummation of the BFKL kernel adopted in the other two approaches. It was maintained in [69, 70] that the diffusion into the ultraviolet, effectively making the coupling weaker, hastens the perturbative convergence for splitting functions, and the kernel near  $\gamma = 0$ , making this additional resummation less necessary. There is no particular obstruction to including this resummation in the approach, it is simply cumbersome. Indeed, in Ref. [70] the effect was checked, and modifications found to be no greater than generic NNLO corrections to the resummation, so it was omitted. (Note that any process where there are two hard scales, sensitive to  $\gamma \approx 0.5$ , or attempted calculation of the hard input for the gluon distribution, sensitive to  $\gamma = 1$ , would find this resummation essential.) The main feature of the resummed splitting functions is a significant dip below the NLO DGLAP results, followed by an eventual rise at very low  $x \simeq 10^{-5}$ . This behaviour drives a qualitative change in the gluon distribution, when implemented in a fit to data.

The combined NLO+NLL splitting and coefficient functions (in the TW approach) have

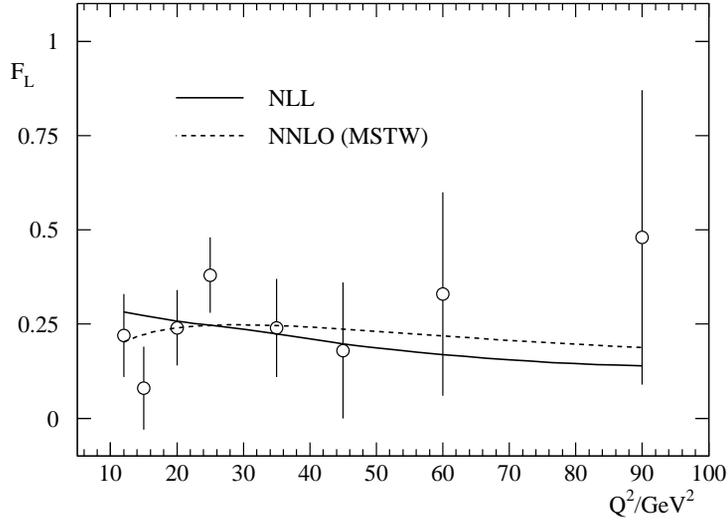


Fig. 10: Recent H1 data on the longitudinal structure function  $F_L$ , together with the NLL resummed prediction from the TW approach, and a recent NNLO result from the MSTW group.

been implemented in a global fit to DIS and related data in the  $\text{DIS}(\chi)$  scheme, thus including small  $x$  resummations in both the massless and massive quark sectors [72]. The overall fit quality was better than a standard NLO fit in the same factorisation scheme, and a similar NLO fit in the more conventional  $\overline{\text{MS}}$  factorisation scheme. The principal reason for this is the dip in the resummed evolution kernels, which allows the gluon distribution to increase at both high and low values of  $x$ . This reduces a tension that exists between the high  $x$  jet data of [86, 87] and the low  $x$  HERA data [17, 88–91]. The gluon distributions arising from the NLL and NLO fits are shown in figure 8, for the starting scale  $Q^2 = 1\text{GeV}^2$  and also for a higher value of  $Q^2$ . One sees that whilst the NLO gluon wants to be negative at low  $x$  and  $Q^2$ , the resummed gluon is positive definite and indeed growing slightly as  $x \rightarrow 0$ . The gluons agree well for higher  $x$  values (where the DGLAP description is expected to dominate), but deviate for  $x \leq 10^{-2}$ . This can therefore be thought of as the value of  $x$  below which resummation starts to become relevant.

The qualitatively different gluon from the resummed fit (together with the decreased evolution kernels w.r.t. the fixed order description) has a number of phenomenological implications:

1. The longitudinal structure function  $F_L$  is sensible at small  $x$  and  $Q^2$  values, where the standard DGLAP description shows a marked instability [92].
2. As a result of the predicted growth of  $F_L$  at small  $x$  the resummed result for the DIS reduced cross-section shows a turnover at high inelasticity  $y$ , in agreement with the HERA

data. This behaviour is not correctly predicted by some fixed order fits.

3. The heavy flavour contribution (from charm and bottom) to  $F_2$  is reduced at higher  $Q^2$  in the resummed approach, due mainly to the decreased evolution, as already noted in a full analysis in the fixed-order expansion at NNLO [93]. Nevertheless, it remains a significant fraction of the total structure function at small  $x$ .

Other resummation approaches should see similar results when confronted with data, given the qualitative (and indeed quantitative) similarities between the splitting functions. It is the decreased evolution with respect to the DGLAP description that drives the qualitative change in the gluon distribution. This is then the source of any quantitative improvement in the description of data, and also the enhanced description of the longitudinal structure function and reduced cross-section.

The resummed prediction for  $F_L$  is shown alongside the recent H1 data [94] in figure 10, and compared with an up-to-date NNLO fixed order result [95]. One sees that the data cannot yet tell apart the predictions, but that they are starting to diverge at low  $x$  and  $Q^2$ , such that data in this range may indeed be sensitive to the differences between resummed and fixed order approaches.

## 2.4 Resummed structure functions: comparison of the ABF and TW approaches

In this section, we present an application of the ABF and TW approaches to the resummed determination of the  $F_2$  and  $F_L$  deep-inelastic structure functions. The corresponding exercise for the CCSS approach has not yet been finalised. A direct comparison of the two approaches is complicated by issues of factorisation scheme dependence: whereas in the ABF approach results may be obtained in any scheme, and in particular the  $\overline{\text{MS}}$  and closely related  $Q_0\text{-}\overline{\text{MS}}$  scheme, in the TW formalism splitting functions and coefficient functions beyond NLO in  $\alpha_S$  are resummed in the  $Q_0\text{-DIS}$  scheme [65, 96], which coincides with the standard DIS scheme at large  $x$  but differs from it at the resummed level; the scheme change needed in order to obtain the coefficient functions from the DIS-scheme ones is performed exactly up to NLO and approximately beyond it. Thus, without a more precise definition of the relation of this scheme to  $\overline{\text{MS}}$ , one cannot compare splitting and coefficient functions, which are factorisation scheme dependent.

A useful compromise is to present the respective results for the ratio of structure function predictions:

$$K_i = \frac{F_i^{NLL}(x, Q^2)}{F_i^{NLO}(x, Q^2)}, \quad (19)$$

where  $i \in 2, L$ , and the  $F_i$  are calculated by convoluting the relevant coefficients with PDFs obtained by perturbative evolution of a common set of partons, defined at a starting scale of  $Q_0^2 = 4\text{GeV}^2$ . The number of flavors is fixed to three, to avoid ambiguities due to heavy quark effects. The initial PDFs are assumed to be fixed (i.e., the same at the unresummed and unresummed level) in the DIS factorization scheme at the scale  $Q_0$ . Of course, in a realistic situation the data are fixed and the PDFs are determined by a fit to the data: hence they are not the same at the resummed and unresummed level (compare Fig. 8 above). However, in the DIS factorization scheme the structure function  $F_2$  is simply proportional to the quark distribution, hence by fixing the PDFs in this scheme one ensures that  $F_2$  is fixed at the starting scale.

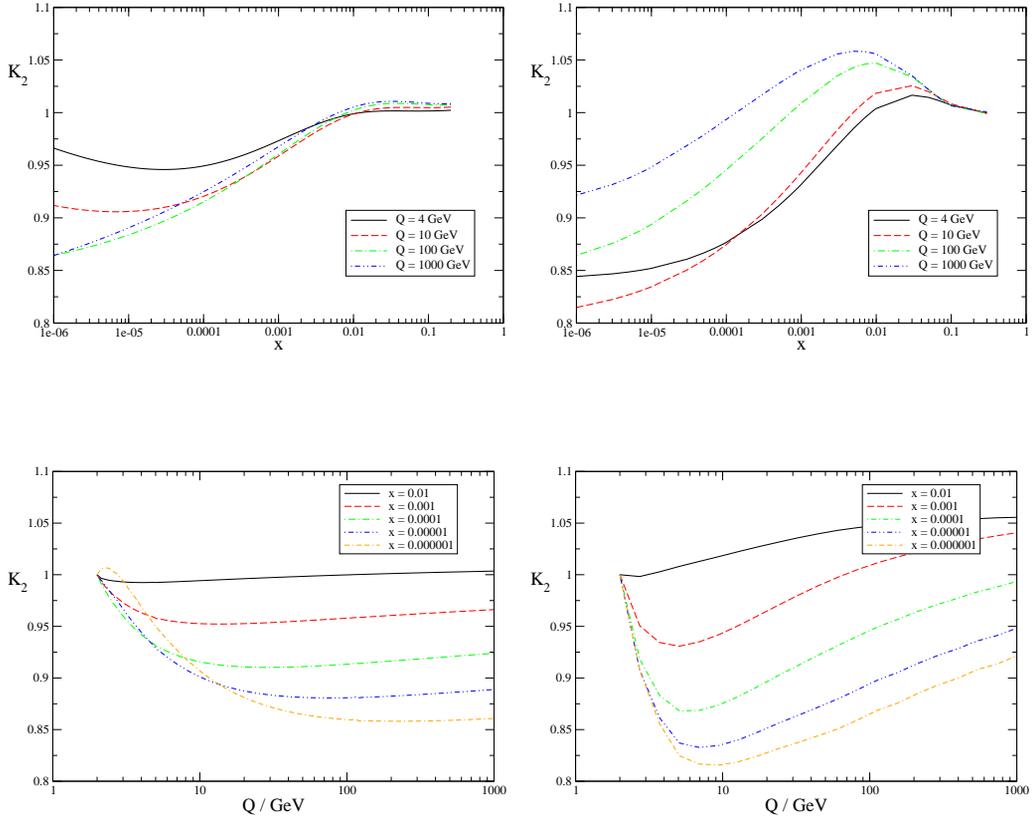


Fig. 11: The ratio  $F_2^{NLL}/F_2^{NLO}$  in the ABF approach (left) and the TW approach (right), using toy PDFs, given in eq. 20, calculated as function of  $x$  at fixed for  $Q^2$  (upper), and as a function of  $Q^2$  at fixed  $x$  (lower).

This starting PDFs are constructed as follows: the quark and gluon distributions are chosen to have the representative form also used in Ref. [58]

$$xg(x) = k_s x S(x) = k_g x^{-0.18} (1-x)^5; \quad xq_v = k_q x^{0.5} (1-x)^4, \quad (20)$$

in the  $\overline{\text{MS}}$  scheme, where  $g(x)$  is the gluon,  $S(x)$  the sea quark distribution, and  $xq_v(x)$  denotes a valence quark distribution. We choose  $k_s = 3$ , and then all other parameters are fixed by momentum and number sum rules. Note that the gluon is the same as that used in the previous comparison of Ref. [73]. The PDFs eq. (20) are then transformed to the DIS factorization scheme [97] using the NLO (unresummed) scheme change at the scale  $Q_0$ . The result is then used as a fixed boundary condition for all (unresummed and resummed, ABF and TW) calculations. In the TW approach, the DIS scheme for unresummed quantities and  $Q_0$ DIS scheme as discussed above is then used throughout. In the ABF approach, the fixed DIS-scheme boundary condition is transformed to the  $Q_0\overline{\text{MS}}$  scheme [58,98] (which at the unresummed level coincides with standard  $\overline{\text{MS}}$ ) by using the unresummed or resummed scheme change function as appropriate, and then all calculations are performed in  $Q_0\overline{\text{MS}}$ . One might hope that most of the residual scheme dependence cancels upon taking the ratio of the NLL and NLO results, at least for schemes that

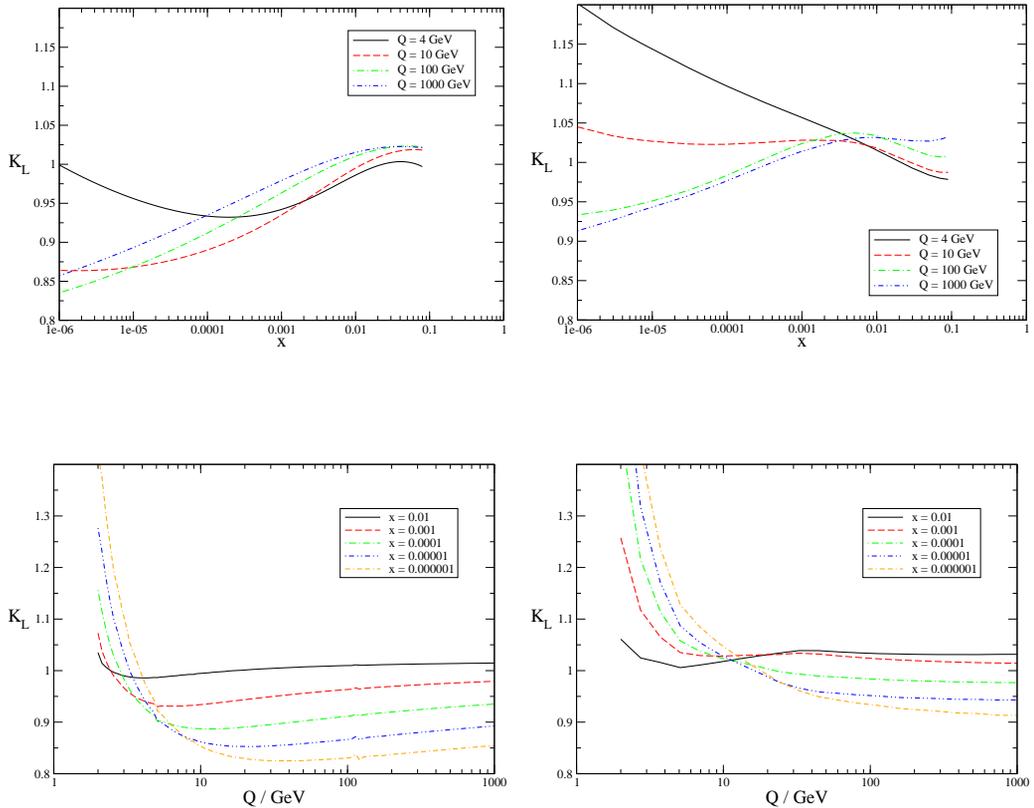


Fig. 12: The ratio  $F_L^{NLL}/F_L^{NLO}$  in the ABF approach (left) and the TW approach (right), using toy PDFs, given in eq. 20, calculated as function of  $x$  at fixed for  $Q^2$  (upper), and as a function of  $Q^2$  at fixed  $x$  (lower).

are well defined and without unphysical singularities.

The results for  $K_2$  and  $K_L$  are shown in figures 11 for  $F_2$  in the ABF and TW procedures respectively and similarly in figures 12 for  $F_L$ . One sees that for  $x$  sufficiently small, and for  $Q$  not too large, the resummed  $F_2$  is consistently lower than its fixed order counterpart in both approaches, due to the decreased evolution of the gluon, and also (in the  $\overline{\text{MS}}$  scheme) due to the fact that resummed coefficient functions are much larger than the NLO ones at small  $x$  and low  $Q^2$ . Similarly the resummed  $F_L$  is larger than the fixed order at low  $Q$  and small enough  $x$ , but falls rapidly as  $Q$  increases. However despite these superficial similarities, the two approaches differ quantitatively in several respects:

- the ABF resummed  $F_2$  matches well to the NLO for  $x \gtrsim 10^{-2}$  at all scales, while the TW  $F_2$  shows a rise around  $x \simeq 10^{-2}$ , which is largest at low  $Q$ . This may be due to the significant differences between resummed and NLO splitting functions at very high  $x$  in fig. 9. A similar mismatch may be seen at  $x \sim 0.1$  in the  $F_L$  K-factor.
- at large scales the ABF resummation stabilises, due to the running of the coupling, so the K-factors becomes rather flat: they grow only logarithmically in  $\ln Q$ . By contrast the TW  $F_2$  K-factor still shows a marked  $Q^2$  dependence. This may be related to the fact that the

TW resummation does not resum the collinear singularities in the BFKL kernel, and to the TW choice (see Sect. 2.3) not to include subtraction of terms induced by the resummation which do not drop at large  $x$ . This choice induces a change in the PDFs at higher  $x$  in the TW approach, which results in effects which persist to higher  $Q^2$  at smaller  $x$ .

- at the initial scale  $Q_0$  the TW resummed  $F_L$  grows much more strongly as  $x$  decreases than the ABF resummed  $F_L$ . This is likely to be due to the different treatment of the coefficient functions: in this respect, the fully consistent treatment of the factorization scheme, the effect of collinear resummation, and the different definitions of what is called resummed NLO used by the two groups all play a part.

## 2.5 Conclusion

The problem of understanding the small  $x$  evolution of structure functions in the domain of  $x$  and  $Q^2$  values of relevance for HERA and LHC physics has by now reached a status where all relevant physical ingredients have been identified, even though not all groups have quite reached the stage at which the formalism can be transformed into a practical tool for a direct connection with the data.

In this report we summarised the status of the three independent approaches to this problem by ABF, CCSS and TW, we discussed the differences in the adopted procedures and finally we gave some recent results. The most complete formalisms are those by ABF and CCSS while the TW approach is less comprehensive but simpler to handle, and thus has been used in fit to data. We recall that, at the level of splitting functions the ABF and CCSS have been compared in ref. [73] and found to be in very good agreement. The singlet splitting function obtained by TW was also compared with ABF and CCSS in ref. [72] and also found to be in reasonable agreement, at least at small  $x$ .

Here we have shown the results of an application to the structure functions  $F_2$  and  $F_L$  of the ABF and TW methods. The same input parton densities at the starting scale  $Q_0$  were adopted by these two groups and the  $K$ -factors for resummed versus fixed NLO perturbative structure functions were calculated using the respective methods. The results obtained are in reasonable qualitative agreement for  $F_2$ , less so for  $F_L$ . Discrepancies may in part be due to the choice of factorization scheme, but our study suggests that the following are also likely to make a quantitative difference: whether or not a resummation of collinear singularities in the BFKL kernel is performed, whether contributions from the resummation which persist at large  $x$  are subtracted and whether the factorization scheme is consistently defined in the same way at resummed and NLO levels.

## 3 Parton saturation and geometric scaling<sup>3</sup>

### 3.1 Introduction<sup>4</sup>

The degrees of freedom involved in hadronic collisions at sufficiently high energy are partons, whose density grows as the energy increases (i.e., when  $x$ , their momentum fraction, decreases). This growth of the number of gluons in the hadronic wave functions is a phenomenon which has

---

<sup>3</sup>Contributing authors: G. Beuf, F. Caola, F. Gelis, L. Motyka, C. Royon, D. Šálek, A. M. Stašo

<sup>4</sup>Contributing authors: F. Gelis, A. M. Stašo

been well established at HERA. One expects however that it should eventually “saturate” when non linear QCD effects start to play a role.

An important feature of partonic interactions is that they involve only partons with comparable rapidities. Consider the interaction between a hadron and some external probe (e.g. a virtual photon in Deep Inelastic Scattering) and consider what happens when one boosts the hadron, increasing its rapidity in successive steps. In the first step, the valence constituents become Lorentz contracted in the longitudinal direction while the time scale of their internal motions is Lorentz dilated. In addition, the boost reveals new vacuum fluctuations coupled to the boosted valence partons. Such fluctuations are not Lorentz contracted in the longitudinal direction, and represent the dynamical degrees of freedom; they are the partons that can interact with the probe. Making an additional step in rapidity would freeze these fluctuations, while making them Lorentz contracted as well. But the additional boost also produces new quantum fluctuations, which become the new dynamical variables. This argument can be repeated, and one arrives at the picture of a high-energy projectile containing a large number of frozen, Lorentz contracted partons (the valence partons, plus all the quantum fluctuations produced in the previous boosts), and partons which have a small rapidity, are not Lorentz contracted and can interact with the probe. This space-time description was developed before the advent of QCD (see for instance [99]; in Bjorken’s lectures [100], one can actually foresee the modern interpretation of parton evolution as a renormalization group evolution).

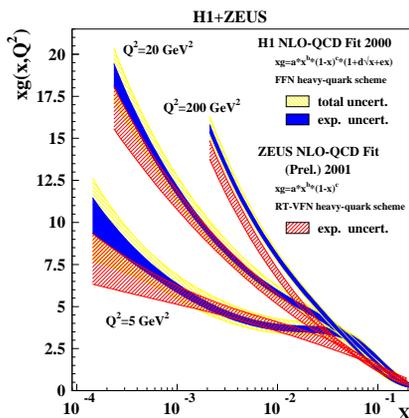


Fig. 13: The gluon structure function in a proton measured at HERA.

This space-time picture, which was deduced from rather general considerations, can now be understood in terms of QCD. In fact, shortly after QCD was established as the theory of strong interaction, quantitative equations were established, describing the phenomenon outlined above [41, 101–105]. In particular, the equation derived by Balitsky, Fadin, Kuraev and Lipatov [41, 101] describes the growth of the non-integrated gluon distribution in a hadron as it is boosted towards higher rapidities. Experimentally, an important increase of the number of gluons at small  $x$  has indeed been observed in the DIS experiments performed at HERA (see Fig. 13), down to  $x \sim 10^{-4}$ . Such a growth raises a problem: if it were to continue to arbitrarily small  $x$ , it would induce an increase of hadronic cross-sections as a power of the center of mass energy, in violation of known unitarity bounds.

However, as noticed by Gribov, Levin and Ryskin in [106], the BFKL equation includes only branching processes that increase the number of gluons ( $g \rightarrow gg$  for instance), but not the recombination processes that could reduce the number of gluons (like  $gg \rightarrow g$ ). While it may be legitimate to neglect the recombination process when the gluon density is small, this cannot remain so at arbitrarily high density: a saturation mechanism of some kind must set in. Treating the partons as ordinary particles, one can get a crude estimate

of the onset of saturation, which occurs at:

$$Q^2 = Q_s^2, \quad \text{with } Q_s^2 \sim \alpha_s(Q_s^2) \frac{xG(x, Q_s^2)}{\pi R^2}. \quad (21)$$

The momentum scale that characterizes this new regime,  $Q_s$ , is called the saturation momentum [107]. Partons with transverse momentum  $Q > Q_s$  are in a dilute regime; those with  $Q < Q_s$  are in the saturated regime. The saturation momentum increases as the gluon density increases. This comes from an increase of the gluon structure function as  $x$  decreases. The increase of the density may also come from the coherent contributions of several nucleons in a nucleus. In large nuclei, one expects  $Q_s^2 \propto A^{1/3}$ , where  $A$  is the number of nucleons in the nucleus.

Note that at saturation, naive perturbation theory breaks down, even though  $\alpha_s(Q_s)$  may be small if  $Q_s$  is large: the saturation regime is a regime of weak coupling, but large density. At saturation, the gluon occupation number is proportional to  $1/\alpha_s$ . In such conditions of large numbers of quanta, classical field approximations become relevant to describe the nuclear wavefunctions.

Once one enters the saturated regime, the evolution of the parton distributions can no longer be described by a linear equation such as the BFKL equation. The color glass condensate formalism (for a review, see [108]), which relies on the separation of the degrees of freedom in a high-energy hadron into frozen partons and dynamical fields, as discussed above, provides the non linear equations that allow us to follow the evolution of the partonic systems from the dilute regime to the dense, saturated, regime. For instance, the correlator  $\text{tr}\langle U^\dagger(\mathbf{x}_\perp)U(\mathbf{y}_\perp) \rangle$  of two Wilson lines –which enters in the discussion of DIS– evolves according to the Balitsky-Kovchegov [109, 110] equation:

$$\begin{aligned} \frac{\partial \text{tr}\langle U^\dagger(\mathbf{x}_\perp)U(\mathbf{y}_\perp) \rangle_x}{\partial \ln(1/x)} &= -\frac{\alpha_s}{2\pi^2} \int_{\mathbf{z}_\perp} \frac{(\mathbf{x}_\perp - \mathbf{y}_\perp)^2}{(\mathbf{x}_\perp - \mathbf{z}_\perp)^2(\mathbf{y}_\perp - \mathbf{z}_\perp)^2} \\ &\times \left[ N_c \text{tr}\langle U^\dagger(\mathbf{x}_\perp)U(\mathbf{y}_\perp) \rangle_x - \text{tr}\langle U^\dagger(\mathbf{x}_\perp)U(\mathbf{z}_\perp) \rangle_x \text{tr}\langle U^\dagger(\mathbf{z}_\perp)U(\mathbf{y}_\perp) \rangle_x \right]. \end{aligned} \quad (22)$$

(This equation reduces to the BFKL equation in the low density limit.)

The geometric scaling phenomenon was first introduced in the context of the dipole picture of the deep inelastic electron-proton scattering [111]. The process of the scattering of the virtual photon on a proton at very small values of  $x$  can be conveniently formulated in the dipole model. In this picture the photon fluctuates into the quark-antiquark pair (dipole) and subsequently interacts with the target. In the small  $x$  regimes these two processes factorize and they can be encoded into the dipole formula for the total  $\gamma^*p$  cross section

$$\sigma_{T,L}(x, Q^2) = \int d^2\mathbf{r} \int dz |\Psi_{T,L}(r, z, Q^2)|^2 \hat{\sigma}(x, r) \quad (23)$$

where  $\Psi_{T,L}$  is the wave function for the photon and  $\hat{\sigma}$  is the dipole cross section.  $r$  is the dipole size and  $z$  is the light-cone fraction of the longitudinal momentum carried by the quark (or antiquark). The photon wave functions  $\Psi$  are known, the dipole cross section can be expressed in terms of the correlator of Wilson lines whose evolution is driven by Eq. (22) :

$$\hat{\sigma}(x, r) = \frac{2}{N_c} \int d^2\mathbf{X} \text{tr} \left\langle 1 - U(\mathbf{X} + \frac{\mathbf{r}}{2}) U^\dagger(\mathbf{X} - \frac{\mathbf{r}}{2}) \right\rangle. \quad (24)$$

Alternatively, it can be modeled or extracted from the data. In the GBW model it was assumed that the dipole cross section has a form

$$\hat{\sigma} = \sigma_0 [1 - \exp(-r^2/R_0(x)^2)] \quad (25)$$

where  $R_0(x) = (x/x_0)^{-\lambda}$  is a saturation radius (its inverse is usually called the saturation scale  $Q_s(x)$ ) and  $\sigma_0$  a normalisation constant. One of the key properties of the model was the dependence on the dipole size and the Bjorken  $x$  through only one combined variable  $r^2 Q_s^2(x)$ . This fact, combined with the property of the dipole formula, allows to reformulate the total cross section as a function of  $Q^2/Q_s^2(x)$  only. This feature is known as the geometric scaling of the total  $\gamma^*p$  cross section. Initially postulated as a property of the GBW model, it was then shown that the experimental data do indeed exhibit the aforementioned regularity in a rather wide range of  $Q^2$  and for small values of Bjorken  $x$ .

Although it is a postulate in the GBW model, this property can be derived from the small- $x$  behavior of the solutions of Eq. (22) [112]: for a wide class of initial conditions, the BK equation drives its solution towards a function that obeys this scaling. Note also that the saturation scale, introduced by hand in the GBW model, is dynamically generated by the non linear evolution described by Eq. (22). This suggested that the regularity seen in the data could be explained by the scaling property of the solutions to the nonlinear equations in the saturated regime - and thus may provide some indirect evidence for gluon saturation.

Nevertheless, several important questions remained. One of them, is the problem of the compatibility of the DGLAP evolution with the property of the geometric scaling. It is known from the global fits that the standard DGLAP evolution works quite well for the description of the of the deep inelastic data even in the very low  $x$  and  $Q^2$  regime. That suggests that the saturation should be confined to the very tight kinematic regime, and it is therefore questionable whether the observed regularity could be attributed to the saturation at all. In the present contribution we discuss several approaches to this problem.

### 3.2 Phenomenology<sup>5</sup>

In order to compare the quality of different scaling laws, it is useful to use a quantity called *quality factor* (QF). It is also used to find the best parameters for a given scaling. In the following, this method is used to compare the scaling results for the proton structure function  $F_2$  and  $F_2^c$ , the deeply virtual Compton scattering, the diffractive structure function, and the vector meson cross section data measured at HERA.

**Quality Factor** Given a set of data points  $(Q^2, x, \sigma = \sigma(Q^2, x))$  and a parametric scaling variable  $\tau = \tau(Q^2, Y, \lambda)$  (with  $Y = \ln 1/x$ ) we want to know whether the cross-section can be parametrised as a function of the variable  $\tau$  only. Since the function of  $\tau$  that describes the data is not known, the *QF* has to be defined independently of the form of that function.

For a set of points  $(u_i, v_i)$ , where  $u_i$ 's are ordered and normalised between 0 and 1, we

---

<sup>5</sup>Contributing authors: C. Royon, D. Šálek

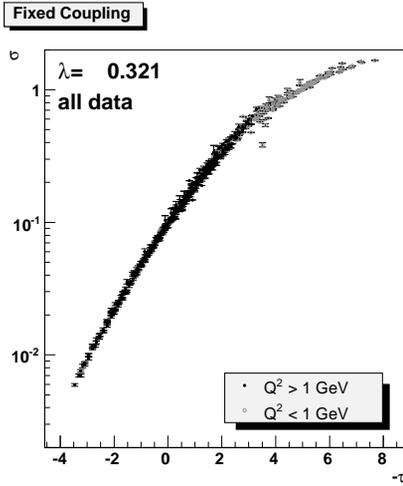


Fig. 14:  $F_2$  data: Scaling curve  $\sigma = \sigma(\tau)$  for “Fixed Coupling”. A  $Q^2 > 1 \text{ GeV}^2$  cut was applied to the data.

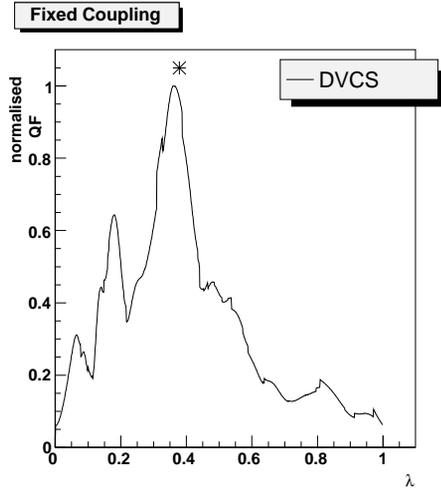


Fig. 15: DVCS data: Quality factor normalised to 1 plotted against the parameter  $\lambda$ . Star denotes the fit result for  $F_2$  data.

introduce  $QF$  as follows [113]

$$QF(\lambda) = \left[ \sum_i \frac{(v_i - v_{i-1})^2}{(u_i - u_{i-1})^2 + \epsilon^2} \right]^{-1}, \quad (26)$$

where  $\epsilon$  is a small constant that prevents the sum from being infinite in case of two points have the same value of  $u$ . According to this definition, the contribution to the sum in (26) is large when two successive points are close in  $u$  and far in  $v$ . Therefore, a set of points lying close to a unique curve is expected to have larger  $QF$  (smaller sum in (26)) compared to a situation where the points are more scattered.

Since the cross-section in data differs by orders of magnitude and  $\tau$  is more or less linear in  $\log(Q^2)$ , we decided to take  $u_i = \tau_i(\lambda)$  and  $v_i = \log(\sigma_i)$ . This ensures that low  $Q^2$  data points contribute to the  $QF$  with a similar weight as higher  $Q^2$  data points.

**Fits to  $F_2$  and DVCS Data** We choose to consider all available data from H1, ZEUS, NMC and E665 experiments [17, 89–91, 114–117] with  $Q^2$  in the range  $[1; 150] \text{ GeV}^2$  and  $x < 0.01^6$ . We exclude the data with  $x > 10^{-2}$  since they are dominated by the valence quark densities, and the formalism of saturation does not apply in this kinematical region. In the same way, the upper  $Q^2$  cut is introduced while the lower  $Q^2$  cut ensures that we stay away from the soft QCD domain. We will show in the following that the data points with  $Q^2 < 1 \text{ GeV}^2$  spoil the fit stability. Two kinds of fits to the scaling laws are performed, either in the full mentioned  $Q^2$  range, or in a tighter  $Q^2$  range  $[3; 150] \text{ GeV}^2$  to ensure that we are in the domain where perturbative QCD applies.

<sup>6</sup>The data in the last ZEUS paper include contributions for  $F_L$  and  $xF_3$  but those can be neglected within the kinematical domain we consider.

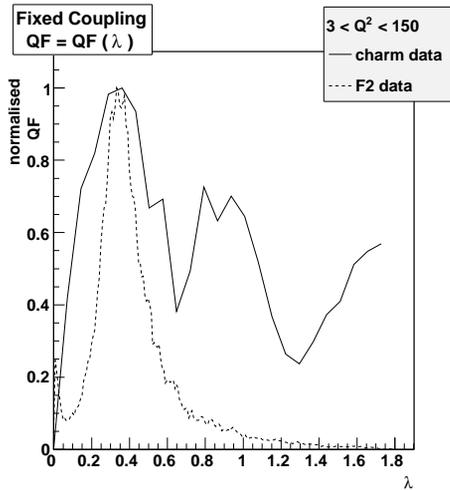


Fig. 16:  $F_2^c$  data: Comparison of the  $\lambda$  parameter for  $F_2$  and  $F_2^c$  data for  $Q^2 > 3 \text{ GeV}^2$ .

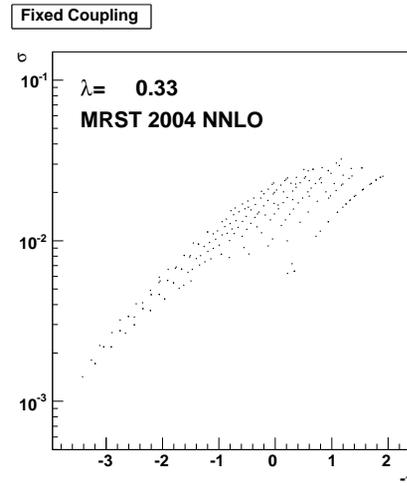


Fig. 17:  $F_2^c$  parametrisation: Scaling curve  $\sigma = \sigma(\tau)$  for fixed coupling using the MRST 2004 NNLO parametrisation for  $\lambda = 0.33$  as obtained in the fit to experimental data. No scaling is observed for  $Q^2 > 3 \text{ GeV}^2$ .

Figure 14 shows the scaling plot for “Fixed Coupling” in the  $Q^2$  range  $[1; 150] \text{ GeV}^2$ , which shows that the lowest  $Q^2$  points in grey have a tendency to lead to worse scaling. The QF values are similar for the “Fixed Coupling”, “Running Coupling I”, and “Running Coupling Ibis” — with a tendency to be slightly better for “Running Coupling Ibis” — and worse for diffusive scaling [118].

The amount of the DVCS data [119, 120] measured by H1 and ZEUS is smaller (34 points for H1 and ZEUS requiring  $x \leq 0.01$  as for  $F_2$  data), therefore the precision on the  $\lambda$  parameter is weaker. The kinematic coverage of the DVCS data covers smaller region in  $x$  and  $Q^2$  than  $F_2$ :  $4 < Q^2 < 25 \text{ GeV}^2$  and  $5 \cdot 10^{-4} < x < 5 \cdot 10^{-3}$ . The DVCS data lead to similar  $\lambda$  values as in the  $F_2$  data (see Fig. 15), showing the consistency of the scalings. The values of the QF show a tendency to favour “Fixed Coupling”, but all different scalings (even “Diffusive Scaling”) lead to reasonable values of QF.

**Implications for Diffraction and Vector Mesons** We used the values of the parameters obtained from the fit to  $F_2$  data to test the various scaling variables on the diffractive cross section and vector meson data [121–123]. We tested both the fixed  $\beta$  scaling behaviour in  $x_{\mathcal{P}}$  and the fixed  $x_{\mathcal{P}}$  scaling behaviour in  $\beta$ . At fixed  $\beta$ , we find a scaling behaviour up to  $\beta = 0.65$ . At fixed  $x_{\mathcal{P}}$ , the scaling behaviour of the diffractive cross section as a function of  $\beta$  and  $Q^2$  is far less obvious. This is not a surprise, as not enough data is available in the genuine small  $\beta$  region. A sign of scaling is however observed for the  $x_{\mathcal{P}} = 0.03$  bin.

Concerning  $\rho$ ,  $J/\Psi$ , and  $\phi$  production [124–126], we found a reasonable scaling behaviour for all tested scaling variables, with the hard scale  $Q^2 + M_V^2$ , borrowed from vector mesons

wave function studies. Surprisingly, the best scaling is for all three vector mesons the “Diffusive scaling”.

**Fits to  $F_2$  and  $F_2^c$  in QCD Parametrisations** First we test the scaling properties using experimental  $F_2^c$  data. The requirements on the kinematical domain remain the same as in the case of  $F_2$  studies. The lower  $Q^2 > 3 \text{ GeV}^2$  cut also allows to remove eventual charm mass effects. We use the charm  $F_2^c$  measurements from the H1 and ZEUS experiments [127–130]. Only 25 data points lie in the desired kinematical region.

Since the statistics in the data is low, the fit results are not precise. Nevertheless, they still lead to clear results that are comparable to  $F_2$  fits. The results are found similar between  $F_2$  and  $F_2^c$  (see Fig. 16). All  $\lambda$  parameters are similar for  $F_2$  and  $F_2^c$  except for “Diffusive Scaling”. As in the case of the  $F_2$  scaling analysis, “Fixed Coupling”, “Running Coupling I” and “Running Coupling II” give similar values of  $QF$ , and “Diffusive Scaling” is disfavoured.

The QCD parametrisations [131–133] of the structure function have been tested using CTEQ, MRST, GRV. The same  $Q^2$  and  $x$  points as in the experimental data were taken into account. Parametrisations of  $F_2$  are able to reproduce the scaling results seen in the experimental data. However, they are not successful in describing the scaling properties in case of  $F_2^c$ . Fig. 17 shows the scaling curve of “Fixed Coupling” in the MRST NNLO 2004 parametrisation of  $F_2^c$  where the value of  $\lambda = 0.33$  is imposed (as seen in the experimental data). The scaling curve is plotted with all the points used in the  $F_2$  study. Therefore the fact that there is not just a single scaling curve in  $F_2^c$  parametrisation is not in direct disagreement with the data — with 25 point only, the curves in parametrisation and data look similar. However the fit values of  $\lambda$  are different.

The CTEQ, MRST or GRV parametrisations are unable to reproduce the scaling properties in  $F_2^c$ . It seems a sea-like intrinsic charm component like the one used in CTEQ 6.6 C4 helps to get results closer to a single scaling curve [134]. Scaling is not present at all in the MRST or GRV parametrisations at low  $Q^2$ .

### 3.3 Geometric scaling and evolution equations with saturation<sup>7</sup>

Let us now recall how scaling properties arise from saturation, as shown in [112], using methods and results from non-linear physics (see [135, 136] for alternative demonstrations). Our discussion, independent of the precise saturation formalism, is valid *e.g.* for the JIMWLK and BK equations (see [108] and references therein), at LL, NLL or even higher order in  $\log(1/x)$ . We will discuss separately the fixed and the running  $\alpha_s$  cases, as running coupling is the main effect which can modify the discussion.

Saturation amounts to add a non-linear damping contribution to the BFKL evolution. One writes formally the evolution equation at LL for the dipole-proton cross section  $\hat{\sigma}$  (23)

$$\partial_Y \hat{\sigma}(Y, L) = \bar{\alpha} \chi(-\partial_L) \hat{\sigma}(Y, L) - \text{non-linear terms in } \hat{\sigma}(Y, L), \quad (27)$$

where  $Y \equiv \log(1/x)$ ,  $L \equiv -\log(r^2 \Lambda_{QCD}^2)$  and  $\chi(\gamma)$  is the characteristic function of the BFKL kernel. The nonlinear damping ensures that, for any  $Y$ ,  $\hat{\sigma}(Y, L)$  grows at most as a power of

---

<sup>7</sup>Contributing author: G. Beuf

$|L|$  for  $L \rightarrow -\infty$  (i.e.  $r \rightarrow +\infty$ ). The color transparency property of the dipole cross section implies  $\hat{\sigma}(Y, L) \propto e^{-L}$  for  $L \rightarrow +\infty$ . Using a double Laplace transform with partial waves  $e^{-\gamma L + \omega Y}$ , the linear part of (27) reduces to the BFKL dispersion relation  $\omega = \bar{\alpha}\chi(\gamma)$ , which gives the partial waves solutions  $e^{-\gamma[L - \bar{\alpha}\chi(\gamma)Y/\gamma]}$ . In the relevant interval  $0 < \gamma < 1$ , the phase velocity  $\lambda(\gamma) = \bar{\alpha}\chi(\gamma)/\gamma$  has one minimum, for the critical value  $\gamma = \gamma_c \simeq 0.63$  which is the solution of  $\chi(\gamma_c) = \gamma_c\chi'(\gamma_c)$ . In the presence of saturation terms in the evolution equation, the wave with  $\gamma = \gamma_c$  is selected dynamically.

In order to understand the dynamics of the problem, let us consider an arbitrary initial condition, at some rapidity  $Y = Y_0$ . With the definition  $\gamma_{eff}(L, Y) \equiv -\partial_L \log(\hat{\sigma}(Y, L))$ ,  $\gamma_{eff}(L, Y_0)$  gives the exponential slope of the initial condition in the vicinity of  $L$ . That vicinity will then propagates for  $Y \geq Y_0$  at a velocity  $\lambda(\gamma_{eff}(L, Y)) = \bar{\alpha}\chi(\gamma_{eff}(L, Y))/\gamma_{eff}(L, Y)$ . One finds easily that, if  $\gamma_{eff}(L, Y_0)$  is a growing function of  $L$ , the regions of smaller velocity will spread during the  $Y$  evolution, and invade the regions of larger velocity. Restricting ourselves to initial conditions verifying the saturation at  $L \rightarrow -\infty$  and the color transparency at  $L \rightarrow +\infty$  as discussed previously, one obtains that  $\gamma_{eff}(L, Y_0)$  goes from 0 at low  $L$  to 1 at large  $L$ . At intermediate  $L$ ,  $\gamma_{eff}(L, Y_0)$  will cross the value  $\gamma_c$ , corresponding to the minimal velocity  $\lambda_c = \lambda(\gamma_c)$ . Hence, one conclude that, as  $Y$  grows, there is a larger and larger domain in  $L$  where  $\gamma_{eff}(L, Y) = \gamma_c$  and thus  $\lambda = \lambda_c$ . In that domain, one has  $\hat{\sigma}(Y, L) \propto e^{-\gamma_c(L - \lambda_c Y)}$ , and hence the geometric scaling  $\hat{\sigma}(Y, L) \equiv f(L - \lambda_c Y) = f(-\log(r^2 Q_s^2(x)))$ , with a saturation scale  $Q_s^2(x) = e^{\lambda_c Y} \Lambda_{QCD}^2 = x^{-\lambda_c} \Lambda_{QCD}^2$ . One finds that the geometric scaling window is limited to  $L < \lambda_c Y + \sqrt{\bar{\alpha}\chi''(\gamma_c)Y/2}$ , and separated from the region still influenced by the initial condition by a cross-over driven by BFKL diffusion. So far, we discussed only scaling properties of the dipole cross section  $\hat{\sigma}$ . As explained in the introduction, they imply similar scaling properties of the virtual photon-proton cross section, with the replacement  $r \mapsto 1/Q$ .

The mechanism of wave selection explained above happens mainly in the linear regime<sup>8</sup>, i.e. for small  $\hat{\sigma}$ , or equivalently  $r$  smaller than  $Q_s^2(x)$ . However, the geometric scaling property stays also valid in the non-linear regime, i.e. for  $r$  larger than  $Q_s^2(x)$ , which is reached after a large enough evolution in  $Y$ . The only, but decisive, role of saturation in the linear domain is to provide the following dynamical boundary condition in the IR to the linear BFKL evolution: when  $\hat{\sigma}$  is large, it should be quite flat ( $\gamma_{eff}(L) \simeq 0$ ). Indeed, one can simulate successfully the impact of saturation on the solution in the linear regime by studying the BFKL evolution in the presence of an absorptive wall [136], set at a  $Y$ -dependent and selfconsistently determined position near the saturation scale.

At NLL and higher order level, the terms different from running coupling ones do not affect the previous discussion. They just change the kernel eigenvalues  $\chi(\gamma)$  and thus shift the selected parameters  $\gamma_c$  and  $\lambda_c$ . On the contrary, going from fixed to running coupling brings important changes. As the mechanism of spreading of smaller velocity regions of the solution towards larger velocity ones is local, one expect that it holds in the running coupling case. But it selects coupling-dependent velocity and shape of the front, the coupling itself being  $L$ -dependent. Hence, the picture is the following. We still have the formation of a specific traveling wave front solution, which progressively loses memory of its initial condition. However, the selected values

<sup>8</sup>We call linear (non-linear) regime the (Y,L) domain where the explicit value of the non-linear terms in (27) is (is not) negligible compared to the value of the linear terms.

of the velocity and shape of the front drift as the front propagate towards larger  $L$  (smaller  $r$ ), due to asymptotic freedom. So far, this running coupling case has been solved analytically [112, 136] only at large  $L$  and large  $Y$ , keeping the relevant geometric scaling variable  $-\log(r^2 Q_s^2(x))$  finite. One finds that the evolution is slower than in the fixed coupling case, as the large  $Y$  behavior of the saturation scale is now  $Q_s^2(x) \sim e^{\sqrt{v_c Y/b}} \Lambda_{QCD}^2$ , with  $b \equiv (33 - 2N_f)/36$  and  $v_c \equiv 2\chi(\gamma_c)/\gamma_c$ . In addition, the geometric scaling window is narrower: asymptotically in  $Y$ , it is expected to hold only for<sup>9</sup>  $L < \sqrt{v_c Y/b} + (|\xi_1|/4) (\chi''(\gamma_c))^{1/3} Y^{1/6} / (2b\gamma_c\chi(\gamma_c))^{1/6}$ . The convergence of the selected front towards this asymptotic solution seems rather slow, which may weaken its phenomenological relevance. The whole theoretical picture is nevertheless consistent with numerical simulations [137, 138]. Both leads to a universal traveling wave front structure of the solution, implying scaling properties also subasymptotically.

In order to do phenomenological studies, one can try to extrapolate to finite  $L$  and  $Y$  the scaling behavior found asymptotically. However, this extrapolation is not unique [139]. There is indeed an infinite family of scaling variables

$$\tau_\delta \equiv \left[ 1 - \left( \frac{v_c Y}{b L^2} \right)^\delta \right] L, \quad (28)$$

parameterized by  $\delta$ , which are different from each other at finite  $L$  and  $Y$  but all converge to the same asymptotic scaling previously mentioned. The parameter  $\delta$  seems quite unconstrained, both from the theory and from the DIS data, as shown in the phenomenological section of the present contribution. We considered as benchmark points in that family two specific choices of  $\delta$ . The choice  $\delta = 1/2$  leads to the only scaling variable of the family which is a genuine geometric scaling variable, *i.e.* is equivalent to a scaling with  $r^2 Q_s^2(x)$ . It is named *running coupling I* in the phenomenological section. The choice  $\delta = 1$  leads to the scaling variable obtained by substitution of the fixed coupling by the running coupling directly in the original fixed coupling geometric scaling variable. It is called *running coupling II*.

Finally, one expects scaling properties in any case from evolution equations with saturation, both in the non-linear regime, and in a scaling window in the linear regime. In the linear regime, the solution still obey the linearized equation, and saturation play only the role of a dynamically generated boundary condition. Hence, geometric scaling there, although generated by saturation, is not a hint against the validity of PDF fits. However, geometric scaling occurs also in the non-linear regime, where the scaling function is no more a solution of the linear BFKL or DGLAP equations.

### 3.4 DGLAP evolution and the saturation boundary conditions<sup>10</sup>

One of the issues that could be studied in the context of the geometric scaling versus DGLAP evolution is the possibility of the different boundary conditions for the DGLAP evolution equations. These boundary conditions would incorporate the saturation effects and posses the scaling property. Typically, in the standard approach, to obtain the solution to the linear DGLAP evolution equations, one imposes the initial conditions onto the parton densities at fixed value of  $Q_0^2$

<sup>9</sup> $\xi_1 \simeq -2.34$  is the rightmost zero of the Airy function.

<sup>10</sup>Contributing author: A. M. Stařto

and then performs the evolution into the region of larger values of  $Q^2$ . However, in the presence of saturation these might not be the correct boundary conditions for DGLAP equations. As mentioned earlier the saturation regime is specified by the critical line, the saturation scale  $Q_s(x)$  which is a function of  $x$  Bjorken and its value increases as the Bjorken  $x$  decreases (or as we go to yet higher energies). In that case it seems legitimate to ask, what is the behavior of the DGLAP solutions when evolved from the saturation boundary  $Q^2 = Q_s^2(x)$  rather than from the fixed scale  $Q^2 = Q_0^2$ . To answer this question we imposed [140] the boundary condition for the gluon density at the saturation scale  $Q^2 = Q_s^2$  which possesses the scaling property namely  $\frac{\alpha_s}{2\pi} xg(x, Q^2 = Q_s^2(x)) = \frac{\alpha_s}{2\pi} r^0 x^{-\lambda}$  (in the fixed coupling case). The solution for the gluon density at small  $x$  (at fixed coupling) which can be derived from solving the DGLAP equations with this boundary is given by

$$\frac{\alpha_s}{2\pi} \frac{xg(x, Q^2)}{Q^2} \sim \frac{\alpha_s}{2\pi} \left( \frac{Q^2}{Q_s^2(x)} \right)^{(\alpha_s/2\pi)\gamma_{gg}(\omega_0)-1} \quad (29)$$

where  $\gamma_{gg}$  is the gluon-gluon DGLAP anomalous dimension. This solution clearly has the geometrical scaling property as it is only a function of  $Q^2/Q_s^2(x)$ . It is interesting to note that there exists a critical value of the exponent  $\lambda$  of the saturation scale which determines the existence of scaling. For example in the double leading logarithmic approximation the scaling is present for rather large values of the exponent  $\lambda \geq 4\alpha_s\pi/3$  whereas there is no scaling for smaller values of  $\lambda$ . The formula shown above is however only approximate, as in the derivation we included only the leading behavior which should be dominant at asymptotically small values of  $x$ . At any finite value of  $x$  the scaling will be mildly violated by the nonleading terms. We checked numerically that this is indeed the case, though the violation was very small. This analysis was extended for the case of the more realistic DGLAP evolution with the running coupling. As expected the presence of the scale violation due to the running coupling will lead to the violation of the scaling. In this case the geometric scaling is only approximate with the solution for the gluon density given by

$$\frac{\alpha_s(Q^2)}{2\pi} \frac{xg(x, Q^2)}{Q^2} \sim \frac{Q_s^2(x)}{Q^2} \left[ 1 + \frac{\alpha_s(Q_s^2(x))}{2\pi b} \ln[Q^2/Q_s^2(x)] \right]^{b\gamma_{gg}(\lambda)-1},$$

with  $b$  being the beta function of the QCD running coupling. The scaling here is present provided we have  $\alpha_s(Q_s(x)) \ln[Q^2/Q_s^2(x)]/(2\pi b) \ll 1$ . Thus the geometric scaling violating term can be factored out.

In summary, this analysis shows that the geometric scaling property can be build into the DGLAP initial conditions, and that the solution to the linear evolution equation which do not include the parton saturation effects can preserve the scaling even in the regime of high  $Q^2$  values, outside the saturation region.

### 3.5 Geometric scaling from DGLAP evolution<sup>11</sup>

From the DGLAP point of view there is another possible explanation for geometric scaling: the scaling behaviour can be generated by the evolution itself, rather than being a preserved boundary condition. In fact, it is possible to show [141] both analytically and numerically that in

---

<sup>11</sup>Contributing author: F. Caola

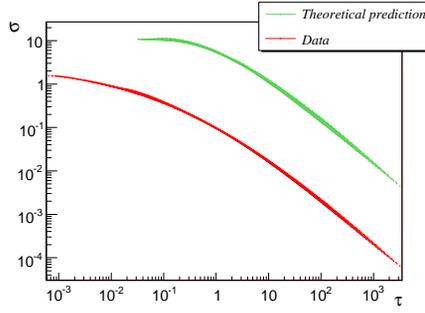


Fig. 18: Scaling plot with  $x < 0.1$ . For the theoretical DGLAP curve, only points with  $Q^2 > 1 \text{ GeV}^2$  were kept. Curves are offset for clarity.

the relevant HERA region approximate geometric scaling is a feature of the DGLAP evolution. In order to see this, one has first to rewrite the DGLAP solution as a function of  $t - \lambda(t, x) \log 1/x$  (“fixed-coupling scaling”) or  $t - \lambda(t, x) \sqrt{\log 1/x}$  (“running-coupling scaling”)<sup>12</sup>. Then from the explicit form of the DGLAP solution it follows that in the relevant kinematic region  $\lambda(t, x)$  is approximately constant, leading to  $\sigma_{DGLAP}(t, x) \approx \sigma_{DGLAP}(t - t_s(x))$ . Hence approximate geometric scaling in the HERA region is a feature of the DGLAP evolution. Interestingly enough, this DGLAP-generated geometric scaling is expected to hold also at large  $Q^2$  and relatively large  $x$  (say  $x \lesssim 0.1$ ), in contrast with the saturation-based geometric scaling which should be a small  $x$ , small (or at least moderate)  $Q^2$  effect.

In order to make more quantitative statements, one can use the quality factor method introduced in Sec. 3.2. As a starting point, one can consider the leading-order small  $x$  DGLAP evolution of a flat boundary condition. At the level of accuracy of geometric scaling, this approximation should be accurate enough in a wide kinematic region, say  $Q^2 \gtrsim 10 \text{ GeV}^2$ ,  $x \lesssim 0.1$  at HERA. Now, a quality-factor analysis shows that in this region the leading-order small  $x$  DGLAP solution has an excellent scaling behaviour, even better than the scaling behaviour observed in HERA data. Also the DGLAP predictions for the geometric slope  $\lambda$  perfectly agree with the phenomenological values: from the DGLAP solution we obtain  $\lambda_{fix}^{DGLAP} = 0.32 \pm 0.05$  (“fixed-coupling” scaling) and  $\lambda_{run}^{DGLAP} = 1.66 \pm 0.34$  (“running-coupling” scaling), to be compared with  $\lambda_{fix}^{exp} = 0.32 \pm 0.06$ ,  $\lambda_{run}^{exp} = 1.62 \pm 0.25$ . Moreover, data exhibit geometric scaling also for larger  $x$ , larger  $Q^2$  (say  $x \lesssim 0.1$  at HERA), as predicted by the DGLAP evolution. All these results are summarized in Fig. 18, where we plot the theoretical and phenomenological<sup>13</sup> reduced cross sections in function of the “fixed-coupling” scaling variable  $\ln \tau = t - \lambda \ln 1/x$ , with  $\lambda = 0.32$ , in the HERA region with the cut  $x < 0.1$ . An analogous plot can be obtained for the “running-coupling” scaling [141]. We interpret these results as striking evidence that for

<sup>12</sup>The labels “fixed-coupling” or “running-coupling” are here a bit misleading. In fact, all the results shown here are obtained with the full running-coupling DGLAP solution. We kept this notation only for comparison with saturation-based approaches.

<sup>13</sup>In fact, in order to make a more flexible analysis, we didn’t use the actual HERA data but a neural network interpolation of world DIS data [142]. As long as one stays in the HERA region the output of the net is totally reliable.

$Q^2 > 10 \text{ GeV}^2$  the geometric scaling seen at HERA is generated by the DGLAP evolution itself, without need of a peculiar saturation ansatz or of a suitable scaling boundary condition.

For  $Q^2 < 10 \text{ GeV}^2$  the leading-order DGLAP solution exhibits violations of geometric scaling at small  $x$ . However, in this region any fixed-order DGLAP calculation fails because it does not resum small  $x$  logarithms. If one consider the DGLAP evolution at the resummed level, geometric scaling reappears quite naturally, both in the "fixed-coupling" and "running-coupling" forms [141]. Hence, small  $x$  resummation extends the region where geometric scaling is expected to values of  $Q^2$  lower than  $10 \text{ GeV}^2$ . However at low  $Q^2$  sizeable higher twist and non perturbative effects can spoil the universal behaviour of the DGLAP solution. In this region hence the HERA scaling could still be generated by some DGLAP evolution, but, differently from the  $Q^2 > 10 \text{ GeV}^2$  region, here there is no strong evidence that this is in fact the case.

### 3.6 Saturation model and higher twists<sup>14</sup>

The QCD description of hard scattering processes within the Operator Product Expansion (OPE) approach leads to the twist expansion of matrix elements of process-dependent composite operators. Contributions of emerging local operators with the increasing twists,  $\tau$ , are suppressed by increasing inverse powers of the hard scale,  $Q^2$ . In DIS, at the lowest order (i.e. when the anomalous dimensions vanish), the twist- $\tau$  contribution to the DIS cross section scales as  $Q^{-\tau}$ . Therefore, at sufficiently large  $Q^2$  it is justified to neglect higher twist effects, and retain only the leading twist-2 contribution. This leads to the standard collinear factorisation approach with universal parton density functions evolving according to the DGLAP evolution equation. It should be kept in mind, however, that the higher twist effects do not vanish completely and that they introduce corrections to theoretical predictions based on the DGLAP approach. Thus, the higher twist corrections may affect the determination of parton density functions. The importance of these corrections depends on the level of precision required and on the kinematic domain. In particular, in the region of very small  $x$  the higher twist effects are expected to be enhanced, so that they may become significant at moderate  $Q^2$ . Thus, it should be useful to obtain reliable estimates of higher twist effects at small  $x$ . In this section we shall present higher twist corrections to  $F_T$ ,  $F_L$  and  $F_2$  structure functions following from the DGLAP improved saturation model [143]. The results presented in this section have been obtained in the course of an ongoing study [144, 145]. The method applied to perform the twist decomposition of the DGLAP improved saturation model is a generalisation of the Mellin space approach proposed in Ref. [146].

A rigorous QCD analysis of the higher twist contributions to DIS at high energies is a complex task. So far it has been performed for the  $q\bar{q}gg$  operators [147], but the evolution of twist 4 purely gluonic operators has not been resolved, — even the proper complete basis of the operators has not been found yet. The collinear evolution is known at all twists, however, for so called *quasi-partonic operators*, for which the twist index is equal to the number of partons in the  $t$ -channel [148]. Such operators should receive the strongest enhancement from the QCD evolution. At the leading logarithmic approximation the collinear evolution of quasi-partonic operators is relatively simple — it is given by pair-wise interactions between the partons in the  $t$ -channel. The interactions are described by the non-forward DGLAP kernel [148]. Within this

---

<sup>14</sup>Contributing author: L. Motyka

formalism, the evolution of four-gluon quasi-partonic operators was investigated in Ref. [149, 150] in the double logarithmic approximation. At small  $x$  the scattering amplitudes are driven by exchange of gluons in the  $t$ -channel, and the quark exchanges are suppressed by powers of  $x$ . Thus we shall focus on the dominant contribution of the multi-gluon exchanges in the  $t$ -channel. In the large  $N_c$ -limit, the dominant singularities of the four gluon operator are those corresponding to states in which gluons get paired into colour singlet states. In other words, the four-gluon operator evolves like a product of two independent gluon densities. In general, for  $1/N_c \rightarrow 0$ , the  $2n$ -gluon (twist- $2n$ ) operator factorizes into the product of  $n$  twist-2 gluon densities. After suitable inclusion of the AGK cutting rules and the symmetry factors of  $1/n!$ , one arrives at the eikonal picture of  $n$ -ladder exchange between the probe and the target. This is to be contrasted with the Balitsky-Kovchegov picture of Pomeron fan diagrams, which was obtained as a result of resummation of the terms enhanced by powers of large  $\ln(1/x)$  rather than by powers of  $\ln Q^2$ .

The eikonal form of the multiple scattering was assumed in the saturation model proposed by Golec-Biernat and Wüsthoff (GBW) [151, 152]. The dipole cross-section given by Eq. 25 has a natural interpretation in terms of a resummation of multiple scattering amplitudes. The scatters are assumed to be independent of each other, and the contribution of  $n$  scatterings is proportional to  $[r^2/R_0^2(x)]^n$ . The connection of the saturation model to the QCD evolution of quasi-partonic operators is further strengthened by the DGLAP improvement of the dipole cross section [143]. In the DGLAP improved saturation model the dipole cross section depends on the collinear gluon density,

$$\hat{\sigma}(x, r) = \sigma_0 \left[ 1 - \exp \left( -\frac{\pi^2 r^2}{N_c \sigma_0} \alpha_s(\mu^2) x g(x, \mu^2) \right) \right], \quad (30)$$

where the scale  $\mu^2$  depends on the dipole size,  $\mu^2 = C/r^2$  for  $C/r^2 > \mu_0^2$ , and  $\mu^2 = \mu_0^2$  for  $C/r^2 < \mu_0^2$ . The gluon density applied has been obtained from the LO DGLAP evolution without quarks, with the input assumed at the scale  $\mu_0^2$ <sup>15</sup>. Clearly, in Eq. (30) one sees an exact matching between the power of  $r^2$  and the power of  $xg(x, \mu^2)$  suggesting a correspondence between the term  $\sim [r^2 \alpha_s(\mu^2) xg(x, \mu^2)]^n$  in the expansion of  $\hat{\sigma}(x, r)$  and the twist- $2n$  contribution to the dipole cross section. Thus, we expect that the saturation model approximately represents higher twist contributions in the deep inelastic scattering generated by the gluonic quasi-partonic operators.

The twist analysis of the DIS cross-section must include a treatment of the quark box that mediates the coupling of the virtual photon,  $\gamma^*$ , to the  $t$ -channel gluons. In the dipole model the  $\gamma^*g \rightarrow q\bar{q}$  amplitude, computed within QCD, is Fourier transformed (w.r.t. the transverse momentum of the quark) to the coordinate representation and appears as the photon wave function, compare Eq. (25). In more detail, one uses the  $\gamma^*g$  amplitude computed within the  $k_T$ -factorisation framework. This amplitude receives contributions from all twists. The twist structure of the quark box is transparent in the space of Mellin moments, and the same is true for the dipole cross-section. Thus we define,

$$\tilde{H}_{T,L}(\gamma, Q^2) = \int_0^1 dz \int_0^\infty dr^2 r^2 |\Psi_{T,L}(r, z, Q^2)|^2 r^{2(\gamma-1)}, \quad (31)$$

<sup>15</sup>In the original DGLAP-improved model [143] a different definition of the scale was adopted,  $\mu^2 = C/r^2 + \mu_0^2$ , but this choice is less convenient for the QCD analysis.

$$\tilde{\sigma}(x, \gamma) = \int_0^\infty dr^2 \hat{\sigma}(x, r^2) r^{2(\gamma-1)}. \quad (32)$$

It then follows from the Parsival formula that,

$$\sigma_{T,L}(x, Q^2) = \int_{\mathcal{C}} \frac{d\gamma}{2\pi i} \tilde{H}_{T,L}(-\gamma, Q^2) \tilde{\sigma}(x, \gamma). \quad (33)$$

For the massless quark case one has  $\tilde{H}_{T,L}(\gamma, Q^2) = \tilde{H}_{T,L}(\gamma) Q^{-2\gamma}$ . The contour of integration,  $\mathcal{C}$ , in Eq. 33 belongs to the fundamental Mellin strip,  $-1 < \text{Re } \gamma < 0$ .

In order to obtain the twist expansion of  $\sigma$ , one extends the contour  $\mathcal{C}$  in the complex  $\gamma$ -plane into a contour  $\mathcal{C}'$  closed from the left-hand side. The Mellin integral in Eq. 33 may be then decomposed into contributions coming from singularities of  $\tilde{H}_{T,L}(-\gamma, Q^2) \tilde{\sigma}(x, \gamma)$ . The function  $\tilde{H}_T(-\gamma)$  ( $\tilde{H}_L(-\gamma)$ ) has simple poles at all negative integer values of  $\gamma$ , except of  $\gamma = -2$  ( $\gamma = -1$ ), where  $\tilde{H}_T$  ( $\tilde{H}_L$ ) is regular. The singularity structure of the dipole cross section,  $\tilde{\sigma}(\gamma)$ , depends on the specific form of  $\hat{\sigma}(x, r^2)$ . For  $\hat{\sigma}(x, r^2)$  used in the GBW model, the  $\tilde{\sigma}(x, \gamma)$  has simple poles at all negative integers  $\gamma$ 's. For the DGLAP improved form of  $\hat{\sigma}$  given by (31),  $\tilde{\sigma}(x, \gamma)$  has cut singularities that extend to the left from  $\gamma = k$  where  $k = -1, -2, \dots$  etc. The leading behaviour of  $\hat{\sigma}$  around a branch point at  $\gamma = k$  is given by  $\sim (\gamma - k)^{p(k)}$ , where the exponent  $p(k)$  is generated by the DGLAP evolution. As the cuts extend to the left from the branch points, the dominant contribution to the cross section at the given twist comes from the vicinity of the corresponding branch point.

The singularity structure of the quark box part  $\tilde{H}_{T,L}(\gamma)$  plays the crucial role in understanding the strength of the subleading twist effects. To see that one expands  $\tilde{H}_{T,L}(\gamma)$  around the singular points,  $\gamma = 1$  and  $\gamma = 2$  (recall that the argument of  $\tilde{H}_{T,L}$  is  $-\gamma$  in the Parsival formula (33)):

$$\tilde{H}_T(\gamma) = \frac{a_T^{(2)}}{\gamma-1} + b_T^{(2)} + \mathcal{O}(\gamma-1), \quad H_L(\gamma) = b_L^{(2)} + \mathcal{O}(\gamma-1), \quad (34)$$

for twist-2, and

$$\tilde{H}_T(\gamma) = b_T^{(4)} + \mathcal{O}(\gamma-2), \quad H_L(\gamma) = \frac{a_L^{(4)}}{\gamma-2} + b_L^{(4)} + \mathcal{O}(\gamma-2), \quad (35)$$

for twist-4. The singular  $1/(\gamma-1)$  and  $1/(\gamma-2)$  terms in (34) and (35) generate an additional enhancement,  $\sim \ln(Q^2)$ , of the corresponding twist-2 and twist-4 contributions to the DIS cross-section. The constant pieces, proportional to  $b_{T,L}^{(2)}$  and  $b_{T,L}^{(4)}$ , produce no new logarithms (thus they are interpreted as the next-to-leading order (NLO) QCD corrections) and the higher terms in the Laurent expansion give yet higher orders in the perturbative expansion of the  $g \rightarrow q$  splitting functions and to the coefficient functions. We summarize this discussion by displaying below the most leading contributions to  $\sigma_{T,L}$  at twist-2 ( $\sigma_{T,L}^{(2)}$ ) and at twist-4 ( $\sigma_{T,L}^{(4)}$ ) obtained in the DGLAP improved saturation model:

$$\sigma_T^{(2)} \sim \frac{a_T^{(2)}}{Q^2} \int_{\mu_0^2}^{Q^2} \frac{dQ'^2}{Q'^2} \alpha_s(Q'^2) xg(x, Q'^2), \quad \sigma_L^{(2)} \sim \frac{b_L^{(2)}}{Q^2} \alpha_s(Q^2) xg(x, Q^2), \quad (36)$$

for twist-2, and

$$\sigma_T^{(4)} \sim \frac{b_T^{(4)}}{Q^4} [\alpha_s(Q^2) x g(x, Q^2)]^2, \quad \sigma_L^{(4)} \sim \frac{a_L^{(4)}}{Q^4} \int_{\mu_0^2}^{Q^2} \frac{dQ'^2}{Q'^2} [\alpha_s(Q'^2) x g(x, Q'^2)]^2, \quad (37)$$

for twist-4. These results imply that the the relative twist-4 correction to  $F_T$  is strongly suppressed w.r.t. the twist-2 contribution, as the subleading twist-4 term in  $F_T$  appears only at the NLO. On the contrary, for  $F_L$ , the leading twist term enters only at the NLO, and the the twist-4 correction enters at the leading order. So, the relative twist-4 effects in  $F_L$  are expected to be enhanced. Note, that both in the case of  $F_T$  and  $F_L$  the twist-4 effects are enhanced w.r.t. the twist-2 contribution by an additional power of the gluon density,  $xg(x, Q^2)$ . For the structure function  $F_2 = F_T + F_L$  we expect small relative corrections from the higher twists because of the opposite sign of coefficients  $a_L^{(4)}$  and  $b_T^{(4)}$ , that leads to cancellations between the twist-4 contributions from  $F_T$  and  $F_L$  at moderate  $Q^2$ . These conclusions about the importance of the higher twist corrections are expected to be quite general, because they follow directly from the twist structure of the quark box and do not depend on the detailed form of the twist-4 gluon distribution.

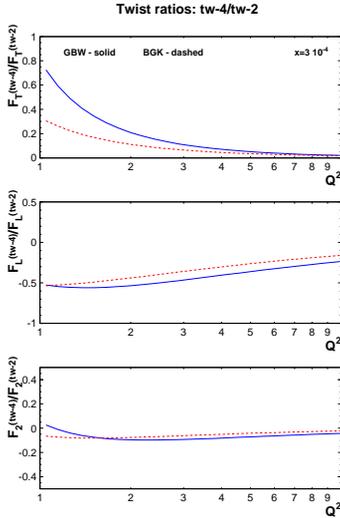


Fig. 19: The ratio of twist-4 to twist-2 components of  $F_T$ ,  $F_L$  and  $F_2$  at  $x = 3 \cdot 10^{-4}$  in the GBW model (continuous lines) and in the DGLAP improved saturation model (dashed lines).

between the twist-4 contributions to  $F_T$  and  $F_L$ , at all  $Q^2$ , down to 1 GeV<sup>2</sup>. Although an effect of this kind was expected, it still remains somewhat surprising that this cancellation works so well. We estimate that, for  $x = 3 \cdot 10^{-4}$ , the twist-4 relative correction to  $F_2$  is 2–4% at  $Q^2 = 10$  GeV<sup>2</sup>, and smaller than 10% for all  $Q^2$  down to 1 GeV<sup>2</sup>. For  $F_L$ , the relative correction is  $\sim 20\%$  at  $Q^2 = 10$  GeV<sup>2</sup>, and strongly increases with the decreasing scale, reaching  $\sim 50\%$  at  $Q^2 = 1$  GeV<sup>2</sup>. It implies that the determination of parton densities from twist-2  $F_2$  data is

We performed [144, 145] an explicit numerical evaluation of the twist-4 corrections to  $F_T$ ,  $F_L$  and  $F_2$  in the DGLAP improved saturation model, and compared the results to results obtained [146] within the GBW model without the DGLAP evolution. The parameters of the DGLAP model were fitted to describe all  $F_2$  data at small  $x$ . In the model we took into account three massless quark flavours and the massive charm quark. The twist analysis, however, has been, so far, performed only for the massless quark contribution. The obtained relative twist-4 corrections to  $F_T$ ,  $F_L$  and  $F_2$  are displayed in Fig. 3.6, as a function of  $Q^2$ , for  $x = 3 \cdot 10^{-4}$ . The continuous curves correspond to the GBW model [146], and the dashed ones have been obtained [144, 145] in the DGLAP improved saturation model. Although there are some quantitative differences between the models, the qualitative picture is quite consistent and confirms the results of the analytic analysis outlined above. Thus, the higher twist corrections are strongest in  $F_L$ , and much weaker in  $F_T$ . In  $F_2$  there occurs a rather fine cancellation

safe even at small  $x$  and moderate  $Q^2$ . On the other hand  $F_L$  at small  $x$  may provide a sensitive probe of higher twist effects and parton saturation.

### 3.7 Conclusions

There are many possible explanations for the scaling properties of HERA data, some of them based on saturation effects and some others based on pure linear evolution. In order to separate between these different explanations, it is fundamental to specify a kinematic window.

In particular, for large enough  $Q^2$  and not too small  $x$  (say  $Q^2 \gtrsim 10 \text{ GeV}^2$  in the HERA region) the observed geometric scaling is determined by the DGLAP evolution, irrespective of the boundary condition. For smaller values of  $Q^2$ , the evolution of parton densities is still linear, but is sensitive to a boundary condition. In an evolution toward smaller  $x$ , like BFKL, this boundary condition is dynamically generated by saturation, and it leads to the geometric scaling window. It is possible to take these effects into account also in a  $Q^2$  evolution, like DGLAP, by imposing as initial condition the same boundary condition. We have seen that, in this case, even the LO DGLAP equation is able to propagate geometric scaling towards larger  $Q^2$ . In that domain, although geometric scaling may arise as saturation effect, the evolution is still linear, and thus compatible with standard PDFs analysis. However, at yet lower  $Q^2$  and  $x$  standard linear evolution is no longer reliable. In particular, for  $Q^2$  smaller than a  $x$  dependent saturation scale  $Q_s(x)$ , the evolution of parton densities becomes fully nonlinear, and this spoils the actual determination of the PDFs. Results from inclusive diffraction and vector meson exclusive production at HERA, and from dA collisions at RHIC all suggest that in the kinematic accessible  $x$  region  $Q_s \sim 1 - 2 \text{ GeV}$ .

In conclusion, we can say that for large enough  $Q^2 \gtrsim 10 \text{ GeV}^2$  geometric scaling is fully compatible with linear DGLAP evolution. For smaller  $Q^2$  the situation becomes more involved. For  $Q^2 \gtrsim 5 \text{ GeV}^2$  the HERA scaling is still compatible with DGLAP, maybe with some small  $x$  resummation or some suitable boundary condition. However, other effects may be relevant in this region. For yet lower  $Q^2$  and  $x$  the linear theory becomes unreliable and saturation could be the right explanation for geometric scaling. Unfortunately at HERA we have too few data for a definitive explanation of geometric scaling in the very small  $x$  region, since many different approaches lead approximatively to the same results and it is very difficult to separate among them. For example, in the low  $x$  region both saturation and perturbative resummations lead to a decrease of the gluon and to geometric scaling. At the LHC, where higher center-of-mass energy is available, the  $x$  region is significantly extended down to very small values. Especially in the fragmentation region the typical values of  $x$  which can be probed can reach down to  $10^{-6}$  for partons with transverse momenta of about few GeV. This fact combined with the very wide rapidity coverage of the main LHC detectors opens up a completely new window for the study of parton saturation, and its relations with geometric scaling and linear evolution will possibly be clarified.

### References

- [1] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, Nucl. Phys. **B427**, 41 (1994).

- [2] S. A. Larin, P. Nogueira, T. van Ritbergen, and J. A. M. Vermaseren, Nucl. Phys. **B492**, 338 (1997). hep-ph/9605317.
- [3] A. Retey and J. A. M. Vermaseren, Nucl. Phys. **B604**, 281 (2001). hep-ph/0007294.
- [4] S. Moch, J. A. M. Vermaseren, and A. Vogt, Nucl. Phys. **B688**, 101 (2004). hep-ph/0403192.
- [5] A. Vogt, S. Moch, and J. A. M. Vermaseren, Nucl. Phys. **B691**, 129 (2004). hep-ph/0404111.
- [6] S. Moch, J. A. M. Vermaseren, and A. Vogt, Phys. Lett. **B606**, 123 (2005). hep-ph/0411112.
- [7] J. A. M. Vermaseren, A. Vogt, and S. Moch, Nucl. Phys. **B724**, 3 (2005). hep-ph/0504242.
- [8] A. Vogt, S. Moch, and J. Vermaseren, Nucl. Phys. Proc. Suppl. **160**, 44 (2006). hep-ph/0608307.
- [9] S. Moch and M. Rogal, Nucl. Phys. **B782**, 51 (2007). 0704.1740.
- [10] S. Moch, M. Rogal, and A. Vogt, Nucl. Phys. **B790**, 317 (2008). 0708.3731.
- [11] A. Vogt, S. Moch, M. Rogal, and J. A. M. Vermaseren, Nucl. Phys. Proc. Suppl. **183**, 155 (2008). 0807.1238.
- [12] S. Moch, M. Rogal, A. Vogt, and J. Vermaseren. In preparation.
- [13] Particle Data Group Collaboration, W. M. Yao *et al.*, J. Phys. **G33**, 1 (2006).
- [14] CCFR/NuTeV Collaboration, U.-K. Yang *et al.*, Phys. Rev. Lett. **86**, 2742 (2001). hep-ex/0009041.
- [15] NuTeV Collaboration, M. Tzanov *et al.*, Phys. Rev. **D74**, 012008 (2006). hep-ex/0509010.
- [16] CHORUS Collaboration, G. Onengut *et al.*, Phys. Lett. **B632**, 65 (2006).
- [17] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C30**, 1 (2003). hep-ex/0304003.
- [18] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C32**, 1 (2003). hep-ex/0307043.
- [19] H1 Collaboration, A. Aktas *et al.*, Phys. Lett. **B634**, 173 (2006). hep-ex/0512060.
- [20] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Lett. **B637**, 210 (2006). hep-ex/0602026.
- [21] M. L. Mangano *et al.*, *Physics at the front-end of a neutrino factory: A quantitative appraisal*. Preprint hep-ph/0105155, 2001.

- [22] J. B. Dainton, M. Klein, P. Newman, E. Perez, and F. Willeke, *JINST* **1**, P10001 (2006). [hep-ex/0603016](#).
- [23] NuTeV Collaboration, G. P. Zeller *et al.*, *Phys. Rev. Lett.* **88**, 091802 (2002). [hep-ex/0110059](#).
- [24] H1 Collaboration, A. Aktas *et al.*, *Phys. Lett.* **B632**, 35 (2006). [hep-ex/0507080](#).
- [25] S. Davidson, S. Forte, P. Gambino, N. Rius, and A. Strumia, *JHEP* **02**, 037 (2002). [hep-ph/0112302](#).
- [26] K. S. McFarland and S.-O. Moch, *Conventional physics explanations for the NuTeV  $\sin^2(\theta(W))$* . Preprint [hep-ph/0306052](#), 2003.
- [27] B. A. Dobrescu and R. K. Ellis, *Phys. Rev.* **D69**, 114014 (2004). [hep-ph/0310154](#).
- [28] S. Kretzer *et al.*, *Phys. Rev. Lett.* **93**, 041802 (2004). [hep-ph/0312322](#).
- [29] A. J. Buras, *Rev. Mod. Phys.* **52**, 199 (1980).
- [30] S. G. Gorishnii, S. A. Larin, L. R. Surguladze, and F. V. Tkachov, *Comput. Phys. Commun.* **55**, 381 (1989).
- [31] S. A. Larin, F. V. Tkachov, and J. A. M. Vermaseren. NIKHEF-H-91-18.
- [32] D. J. Broadhurst, A. L. Kataev, and C. J. Maxwell, *Phys. Lett.* **B590**, 76 (2004). [hep-ph/0403037](#).
- [33] W. L. van Neerven and A. Vogt, *Nucl. Phys.* **B603**, 42 (2001). [hep-ph/0103123](#).
- [34] E. A. Paschos and L. Wolfenstein, *Phys. Rev.* **D7**, 91 (1973).
- [35] S. Moch, J. A. M. Vermaseren, and A. Vogt, *Nucl. Phys.* **B726**, 317 (2005). [hep-ph/0506288](#).
- [36] S. Catani, D. de Florian, G. Rodrigo, and W. Vogelsang, *Phys. Rev. Lett.* **93**, 152003 (2004). [hep-ph/0404240](#).
- [37] H. L. Lai *et al.*, *JHEP* **04**, 089 (2007). [hep-ph/0702268](#).
- [38] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, *Parton Distributions for the LHC*. Preprint 0706.0456, 2007.
- [39] A. Vogt (2007). [arXiv:0707.4106](#).
- [40] V. S. Fadin, E. A. Kuraev, and L. N. Lipatov, *Phys. Lett.* **B60**, 50 (1975).
- [41] I. I. Balitsky and L. N. Lipatov, *Sov. J. Nucl. Phys.* **28**, 822 (1978).
- [42] V. S. Fadin and L. N. Lipatov, *Phys. Lett.* **B429**, 127 (1998). [hep-ph/9802290](#).

- [43] G. Camici and M. Ciafaloni, Phys. Lett. **B412**, 396 (1997). hep-ph/9707390.
- [44] M. Ciafaloni and G. Camici, Phys. Lett. **B430**, 349 (1998). hep-ph/9803389.
- [45] S. Marzani, R. D. Ball, P. Falgari, and S. Forte, Nucl. Phys. **B783**, 143 (2007). 0704.2404.
- [46] M. Ciafaloni, D. Colferai, D. Colferai, G. P. Salam, and A. M. Stasto, Phys. Lett. **B576**, 143 (2003). hep-ph/0305254.
- [47] M. Ciafaloni, D. Colferai, G. P. Salam, and A. M. Stasto, Phys. Rev. **D68**, 114003 (2003). hep-ph/0307188.
- [48] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).
- [49] S. Catani and F. Hautmann, Nucl. Phys. **B427**, 475 (1994). hep-ph/9405388.
- [50] A. Bialas, H. Navelet, and R. B. Peschanski, Nucl. Phys. **B603**, 218 (2001). hep-ph/0101179.
- [51] C. D. White, R. B. Peschanski, and R. S. Thorne, Phys. Lett. **B639**, 652 (2006). hep-ph/0606169.
- [52] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B575**, 313 (2000). hep-ph/9911273.
- [53] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B621**, 359 (2002). hep-ph/0109178.
- [54] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B674**, 459 (2003). hep-ph/0306156.
- [55] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B742**, 1 (2006). hep-ph/0512237.
- [56] R. D. Ball and S. Forte, Nucl. Phys. **B742**, 158 (2006). hep-ph/0601049.
- [57] R. D. Ball, Nucl. Phys. **B796**, 137 (2008). 0708.1277.
- [58] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B799**, 199 (2008). 0802.0032.
- [59] G. Altarelli, R. D. Ball, and S. Forte, *Structure Function Resummation in small- $x$  QCD*. Preprint 0802.0968, 2007.
- [60] G. P. Salam, JHEP **07**, 019 (1998). hep-ph/9806482.
- [61] M. Ciafaloni and D. Colferai, Phys. Lett. **B452**, 372 (1999). hep-ph/9812366.
- [62] M. Ciafaloni, D. Colferai, and G. P. Salam, Phys. Rev. **D60**, 114036 (1999). hep-ph/9905566.
- [63] M. Ciafaloni, D. Colferai, and G. P. Salam, JHEP **07**, 054 (2000). hep-ph/0007240.

- [64] M. Ciafaloni and D. Colferai, JHEP **09**, 069 (2005). hep-ph/0507106.
- [65] M. Ciafaloni, D. Colferai, G. P. Salam, and A. M. Stasto, Phys. Lett. **B635**, 320 (2006). hep-ph/0601200.
- [66] M. Ciafaloni, D. Colferai, G. P. Salam, and A. M. Stasto, JHEP **08**, 046 (2007). 0707.1453.
- [67] R. S. Thorne, Phys. Rev. **D60**, 054031 (1999). hep-ph/9901331.
- [68] R. S. Thorne, Nucl. Phys. Proc. Suppl. **79**, 210 (1999). hep-ph/9906323.
- [69] R. S. Thorne, Phys. Lett. **B474**, 372 (2000). hep-ph/9912284.
- [70] R. S. Thorne, Phys. Rev. **D64**, 074005 (2001). hep-ph/0103210.
- [71] C. D. White and R. S. Thorne, Phys. Rev. **D74**, 014002 (2006). hep-ph/0603030.
- [72] C. D. White and R. S. Thorne, Phys. Rev. **D75**, 034005 (2007). hep-ph/0611204.
- [73] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.
- [74] R. D. Ball and S. Forte, Phys. Lett. **B465**, 271 (1999). hep-ph/9906222.
- [75] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. **B599**, 383 (2001). hep-ph/0011270.
- [76] G. Altarelli, R. D. Ball, and S. Forte, *An improved splitting function for small  $x$  evolution*. Preprint hep-ph/0310016, 2003.
- [77] G. Altarelli, R. D. Ball, and S. Forte, Nucl. Phys. Proc. Suppl. **135**, 163 (2004). hep-ph/0407153.
- [78] R. D. Ball and S. Forte, Phys. Lett. **B405**, 317 (1997). hep-ph/9703417.
- [79] L. Lipatov, Sov. Phys. JETP **5**, 5 (1986).
- [80] R. D. Ball and R. K. Ellis, JHEP **05**, 053 (2001). hep-ph/0101199.
- [81] B. Andersson, G. Gustafson, H. Kharraziha, and J. Samuelsson, Z. Phys. **C71**, 613 (1996).
- [82] J. Kwiecinski, A. D. Martin, and P. J. Sutton, Z. Phys. **C71**, 585 (1996). hep-ph/9602320.
- [83] J. Kwiecinski, A. D. Martin, and A. M. Stasto, Phys. Rev. **D56**, 3991 (1997). hep-ph/9703445.
- [84] J. C. Collins and J. Kwiecinski, Nucl. Phys. **B316**, 307 (1989).

- [85] C. D. White and R. S. Thorne, *Eur. Phys. J.* **C45**, 179 (2006). hep-ph/0507244.
- [86] D0 Collaboration, B. Abbott *et al.*, *Phys. Rev. Lett.* **86**, 1707 (2001).  
hep-ex/0011036.
- [87] CDF Collaboration, T. Affolder *et al.*, *Phys. Rev.* **D64**, 032001 (2001).  
hep-ph/0102074.
- [88] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C19**, 269 (2001). hep-ex/0012052.
- [89] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C21**, 33 (2001). hep-ex/0012053.
- [90] ZEUS Collaboration, J. Breitweg *et al.*, *Eur. Phys. J.* **C7**, 609 (1999).  
hep-ex/9809005.
- [91] ZEUS Collaboration, S. Chekanov *et al.*, *Eur. Phys. J.* **C21**, 443 (2001).  
hep-ex/0105090.
- [92] A. D. Martin, W. J. Stirling, and R. S. Thorne, *Phys. Lett.* **B635**, 305 (2006).  
hep-ph/0601247.
- [93] R. S. Thorne, *Phys. Rev.* **D73**, 054019 (2006). hep-ph/0601245.
- [94] H1 Collaboration, F. D. Aaron *et al.*, *Phys. Lett.* **B665**, 139 (2008). 0805.2809.
- [95] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Phys. Lett.* **B652**, 292 (2007).  
0706.0459.
- [96] M. Ciafaloni, *Phys. Lett.* **B356**, 74 (1995). hep-ph/9507307.
- [97] M. Diemoz, F. Ferroni, E. Longo, and G. Martinelli, *Z. Phys.* **C39**, 21 (1988).
- [98] R. D. Ball and S. Forte, *Phys. Lett.* **B359**, 362 (1995). hep-ph/9507321.
- [99] P. F. R. *Photon-Hadron Interactions*. Benjamin, New York, 1972.
- [100] Bjorken, J.D., *Lecture Notes in Physics*, **56**, Springer, Berlin (1976).
- [101] Kuraev, E.A. and Lipatov, L.N. and Fadin, V.S., *Sov. Phys. JETP* **45**, 199 (1977).
- [102] N. G. V. and N. L. L. *Sov. J. Nucl. Phys.* **15**, 438 (1972).
- [103] N. G. V. and N. L. L. *Sov. J. Nucl. Phys.* **15**, 675 (1972).
- [104] G. Altarelli and G. Parisi, *Nucl. Phys.* **B 126**, 298 (1977).
- [105] L. D. Yu. *Sov. Phys. JETP* **46**, 641 (1977).
- [106] Gribov, L.V. and Levin, E.M. and Ryskin, M.G., *Phys. Rept.* **100**, 1 (1983).
- [107] Mueller, A.H., *Nucl. Phys.* **B558**, 285 (1999).

- [108] E. Iancu and R. Venugopalan, *The color glass condensate and high energy scattering in QCD*. Preprint hep-ph/0303204, 2003.
- [109] Balitsky, I., Nucl. Phys. **B463**, 99 (1996).
- [110] Kovchegov, Yu.V., Phys. Rev. **D61**, 074018 (2000).
- [111] A. M. Stasto, K. J. Golec-Biernat, and J. Kwiecinski, Phys. Rev. Lett. **86**, 596 (2001). hep-ph/0007192.
- [112] S. Munier and R. Peschanski, Phys. Rev. **D69**, 034008 (2004). hep-ph/0310357.
- [113] F. Gelis, R. B. Peschanski, G. Soyez, and L. Schoeffel, Phys. Lett. **B647**, 376 (2007). hep-ph/0610435.
- [114] ZEUS Collaboration, J. Breitweg *et al.*, Phys. Lett. **B487**, 273 (2000). hep-ex/0006013.
- [115] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Rev. **D70**, 052001 (2004). hep-ex/0401003.
- [116] New Muon Collaboration Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B483**, 3 (1997). hep-ph/9610231.
- [117] E665 Collaboration, M. R. Adams *et al.*, Phys. Rev. **D54**, 3006 (1996).
- [118] G. Beuf, R. Peschanski, C. Royon, and D. Salek, *Systematic Analysis of Scaling Properties in Deep Inelastic Scattering*. Preprint arXiv:0803.2186 [hep-ph], 2008.
- [119] H1 Collaboration, F. D. Aaron *et al.*, Phys. Lett. **B659**, 796 (2008). 0709.4114.
- [120] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C44**, 1 (2005). hep-ex/0505061.
- [121] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C48**, 715 (2006). hep-ex/0606004.
- [122] ZEUS Collaboration, S. Chekanov *et al.*, Nucl. Phys. **B713**, 3 (2005). hep-ex/0501060.
- [123] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C38**, 43 (2004). hep-ex/0408009.
- [124] ZEUS Collaboration, S. Chekanov *et al.*, Nucl. Phys. **B718**, 3 (2005). hep-ex/0504010.
- [125] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C46**, 585 (2006). hep-ex/0510016.
- [126] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C13**, 371 (2000). hep-ex/9902019.
- [127] H1 Collaboration, C. Adloff *et al.*, Z. Phys. **C72**, 593 (1996). hep-ex/9607012.
- [128] H1 Collaboration, C. Adloff *et al.*, Phys. Lett. **B528**, 199 (2002). hep-ex/0108039.

- [129] ZEUS Collaboration, J. Breitweg *et al.*, Phys. Lett. **B407**, 402 (1997).  
hep-ex/9706009.
- [130] European Muon Collaboration, J. J. Aubert *et al.*, Nucl. Phys. **B213**, 31 (1983).
- [131] P. M. Nadolsky *et al.*, Phys. Rev. **D78**, 013004 (2008). 0802.0007.
- [132] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Phys. Lett.  
**B604**, 61 (2004). hep-ph/0410230.
- [133] M. Gluck, E. Reya, and A. Vogt, Eur. Phys. J. **C5**, 461 (1998). hep-ph/9806404.
- [134] R. C. Beuf, G. and D. Salek, to appear.
- [135] E. Iancu, K. Itakura, and L. McLerran, Nucl. Phys. **A708**, 327 (2002).  
hep-ph/0203137.
- [136] A. H. Mueller and D. N. Triantafyllopoulos, Nucl. Phys. **B640**, 331 (2002).  
hep-ph/0205167.
- [137] E. Gardi, J. Kuokkanen, K. Rummukainen, and H. Weigert, Nucl. Phys.  
**A784**, 282 (2007). hep-ph/0609087.
- [138] J. L. Albacete and Y. V. Kovchegov, Phys. Rev. **D75**, 125021 (2007).  
arXiv:0704.0612 [hep-ph].
- [139] G. Beuf, *An alternative scaling solution for high-energy qcd saturation with running coupling*. Preprint arXiv:0803.2167 [hep-ph], 2008.
- [140] J. Kwiecinski and A. M. Stasto, Phys. Rev. **D66**, 014013 (2002). hep-ph/0203030.
- [141] F. Caola and S. Forte, Phys. Rev. Lett. **101**, 022001 (2008). 0802.1878.
- [142] NNPDF Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo,  
JHEP **03**, 080 (2005). hep-ph/0501067.
- [143] J. Bartels, K. J. Golec-Biernat, and H. Kowalski, Phys. Rev. **D66**, 014001 (2002).  
hep-ph/0203258.
- [144] K. G.-B. J. Bartels and L. Motyka, *in preparation*.
- [145] L. Motyka, *Higher twists from the saturation model*.  
Talk at the 4th HERA and the LHC workshop, CERN, 26–30 May 2008,  
<http://indico.cern.ch/conferenceDisplay.py?confId=27458>.
- [146] J. Bartels, K. J. Golec-Biernat, and K. Peters, Eur. Phys. J. **C17**, 121 (2000).  
hep-ph/0003042.
- [147] R. K. Ellis, W. Furmanski, and R. Petronzio, Nucl. Phys. **B212**, 29 (1983).

- [148] A. P. Bukhvostov, G. V. Frolov, L. N. Lipatov, and E. A. Kuraev, Nucl. Phys. **B258**, 601 (1985).
- [149] J. Bartels and M. G. Ryskin, Z. Phys. **C60**, 751 (1993).
- [150] J. Bartels and M. G. Ryskin, Z. Phys. **C62**, 425 (1994).
- [151] K. J. Golec-Biernat and M. Wusthoff, Phys. Rev. **D59**, 014017 (1999).  
hep-ph/9807513.
- [152] K. J. Golec-Biernat and M. Wusthoff, Phys. Rev. **D60**, 114023 (1999).  
hep-ph/9903358.

# Benchmarking of parton distributions and their uncertainties

*R. D. Ball, L. Del Debbio, J. Feltesse, S. Forte, A. Glazov, A. Guffanti, J. I. Latorre, A. Piccione, V. Radescu, J. Rojo, R. S. Thorne, M. Ubiali, G. Watt*

## 1 Introduction

The proper treatment of uncertainties associated to the fit of Parton Distribution Functions (PDF) has become a subject of great interest in the last few years. A simple way of understanding differences between available approaches to parton fits is to fix some hypothesis (say, experimental data, QCD parameters, input parameterizations, error treatment), and check what is the effect of the remaining assumptions. Such studies were previously done in the framework of the first HERA–LHC workshop [1].

In the following we will discuss three benchmark fits. The first one is presented in Sect. 2. It is based on the H12000 parton fit [2], and it compares a new version of this fit, in which uncertainty bands are determined [3,4] using a Monte Carlo method, to the reference fit, where uncertainty bands are obtained using the standard Hessian method. The main motivation of this benchmark is to study the impact of possible non-Gaussian behaviour of the data and, more generally, the dependence on the error treatment.

The second benchmark is presented in Sect. 3. It is based on the study performed by S. Alekhin and R. Thorne in Ref. [1], which compared the fits by their respective groups to a common reduced set of data with common assumptions, and also to their respective reference (global) fits. This comparison is extended here in two ways. First, the comparison is extended to include an NNPDF fit to the same reduced set of data with the same assumptions, and the NNPDF1.0 reference fit [5]. Second, results are also compared to a fit based on the recent MSTW 2008 [6, 7] analysis. As in the Thorne benchmark fit, this uses slightly different data sets and assumptions; it is furthermore modified to use the same input parameterization and improved treatment of uncertainties as MSTW. The main purpose of these comparisons is to answer the questions (a) to which extent fit results from various groups obtained using different methodologies still differ from each other when common or similar assumptions and a common or similar reduced dataset are used and (b) how the fits to the reduced dataset by each group compare to the fit to the full dataset.

The third benchmark, discussed in Sect. 4, is a further elaboration on the benchmark presented in Sect. 2, extended to include the NNPDF fit, which also uses a Monte Carlo approach. The main purpose of this benchmark is to compare two fits (H1 and NNPDF) which have the same error treatment but different parton parameterizations. The inclusion in this benchmark of the NNPDF fit is also interesting because it allows a comparison of a fit based on a very consistent set of data coming from the H1 collaboration only, to fits which include all DIS data sets, which are less compatible than the H1 sets alone.

### 1.1 Settings for the H1 benchmark

This analysis is based on all the DIS inclusive data by the H1 collaboration from the HERA-I run. A kinematic cut of  $Q^2 > 3.5 \text{ GeV}^2$  is applied to avoid any higher twist effect. The data

points used in the analysis are summarized in Table 1 and Fig. 1.

Data Set	Data points	Observable	Ref.
H197mb	35	$\tilde{\sigma}^{NC,+}$	[8]
H197lowQ2	80	$\tilde{\sigma}^{NC,+}$	[8]
H197NC	130	$\tilde{\sigma}^{NC,+}$	[9]
H197CC	25	$\tilde{\sigma}^{CC,+}$	[9]
H199NC	126	$\tilde{\sigma}^{NC,-}$	[10]
H199CC	28	$\tilde{\sigma}^{CC,-}$	[10]
H199NChy	13	$\tilde{\sigma}^{NC,-}$	[10]
H100NC	147	$\tilde{\sigma}^{NC,+}$	[2]
H100CC	28	$\tilde{\sigma}^{CC,+}$	[2]
Total	612		

Table 1: Data points used in the H1 benchmark after kinematic cuts of  $Q^2 > 3.5 \text{ GeV}^2$ .

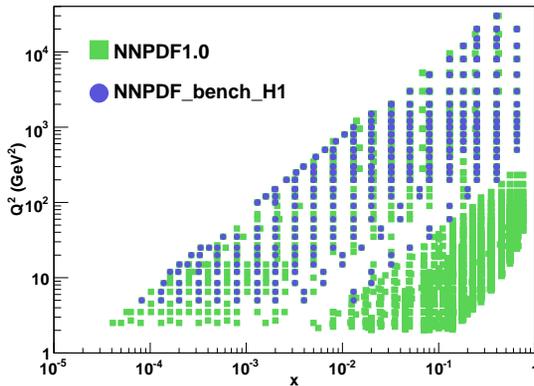


Fig. 1: The data used in the H1 benchmark and in the NNPDF reference fit.

The theoretical assumptions are:

- NLO perturbative QCD in the  $\overline{\text{MS}}$  renormalization and factorization scheme;
- zero-mass variable flavour number scheme with quark masses  $m_c = 1.4 \text{ GeV}$  and  $m_b = 4.5 \text{ GeV}$ ;
- the strong coupling fixed to  $\alpha_s(M_Z) = 0.1185$ ;
- momentum and valence sum rules enforced;
- starting scale for the evolution at  $Q_0^2 = 4 \text{ GeV}^2$ ;
- strange contribution fixed as

$$s(x, Q_0^2) = \bar{s}(x, Q_0^2) = f_s \bar{D}(x, Q_0^2) = \frac{f_s}{1 - f_s} \bar{d}(x, Q_0^2), \quad (1)$$

with  $U = u + c$  and  $D = d + s + b$  and with  $f_s = 0.33$ ;

- charm contribution fixed as

$$c(x, Q_0^2) = \bar{c}(x, Q_0^2) = f_c \bar{U}(x, Q_0^2) = \frac{f_c}{1 - f_c} \bar{u}(x, Q_0^2), \quad (2)$$

with  $f_c = 0.15$ ;

- five independent PDFs: gluon and  $U, D, \bar{U}, \bar{D}$  (see definition above);
- iterated solution for evolution (see, e.g. [11], Sect. 1.3).

Both the H1 and NNPDF methodologies are based on

- Monte Carlo method to determine uncertainties. This method will be discussed in detail in Sect. 2.2 below.

They differ in the way PDFs are parameterized:

- H1 parameterizes PDFs as

$$\begin{aligned} xg(x, Q_0^2) &= A_g x^{B_g} (1-x)^{C_g} [1 + D_g x], \\ xU(x, Q_0^2) &= A_U x^{B_U} (1-x)^{C_U} [1 + D_U x + F_U x^3], \\ xD(x, Q_0^2) &= A_D x^{B_D} (1-x)^{C_D} [1 + D_D x], \\ x\bar{U}(x, Q_0^2) &= A_{\bar{U}} x^{B_{\bar{U}}} (1-x)^{C_{\bar{U}}}, \\ x\bar{D}(x, Q_0^2) &= A_{\bar{D}} x^{B_{\bar{D}}} (1-x)^{C_{\bar{D}}}, \end{aligned} \quad (3)$$

(4)

which yields 10 free parameters after sum rules are imposed;

- NNPDF parameterizes PDFs with a 2-5-3-1 neural network, which implies 185 free parameters to be fitted.

Because of the large number of parameters, the minimum of the NNPDF fit is determined using the stopping criterion discussed in Sect. 3.2 below, while the minimum of the H1 fit is determined as the standard minimum  $\chi^2$  (or maximum likelihood) point of parameter space.

## 1.2 Settings for the HERA–LHC benchmark

This benchmark was first presented in Ref. [1], where its settings were defined. In order to have a conservative ensemble of experimental data and observables, only structure function DIS data are used. Large kinematic cuts are applied to avoid any higher twist effect. The data points used in the Alekhin analysis are summarized in Table 2 and Fig. 2.

The theoretical assumptions are:

- NLO perturbative QCD in the  $\overline{\text{MS}}$  renormalization and factorization scheme;
- zero-mass variable flavour number scheme with quark masses  $m_c = 1.5 \text{ GeV}$  and  $m_b = 4.5 \text{ GeV}$ ;
- $\alpha_s(M_Z)$  fitted: the best-fit values are  $0.1110 \pm 0.0012$  (Alekhin) and  $0.1132 \pm 0.0015$  (Thorne);
- momentum and valence sum rules imposed;

Data Set	Data points	Observable	Ref.
ZEUS97	206	$F_2^p$	[12]
H1lowx97	77	$F_2^p$	[8]
NMC	95	$F_2^p$	[13]
NMC_pd	73	$F_2^d/F_2^p$	[14]
BCDMS	322	$F_2^p$	[15]
Total	773		

Table 2: Data points used in the HERA–LHC benchmark after kinematic cuts of  $Q^2 > 9 \text{ GeV}^2$  and  $W^2 > 15 \text{ GeV}^2$  are applied.

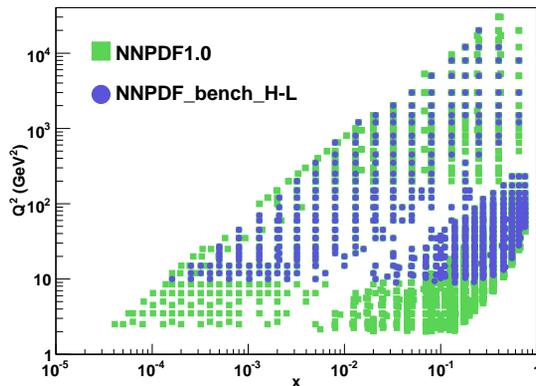


Fig. 2: The data used in the HERA–LHC benchmark and in the NNPDF reference fit.

- starting scale for evolution  $Q_0^2 = 1 \text{ GeV}^2$ ;
- four independent input PDFs ( $u$  and  $d$  valence, the sea and the gluon);
- no light sea asymmetry:  $\bar{u} = \bar{d}$ ;
- no independent strange PDF:

$$s(x, Q_0^2) + \bar{s}(x, Q_0^2) = 0.5(\bar{u}(x, Q_0^2) + \bar{d}(x, Q_0^2)); \quad (5)$$

- iterated solution of evolution equations;

The NNPDF analysis presented here is based on the same data set and theoretical assumptions, the only difference being that the strong coupling is fixed to  $\alpha_s(M_Z) = 0.112$ , i.e. the average of the fitted values of S. Alekhin and R. Thorne.

The Thorne benchmark used somewhat different data sets and assumptions. Namely:

- A somewhat different dataset is used, as displayed in Table 3. This differs from the dataset of Table 2 and Figure 2 because the NMC and BCDMS fixed-target data on  $F_2^p$  used are averaged over different beam energies, and also, HERA reduced cross sections rather than structure function data are used, resulting in an additional nine H1 points. Note that the Thorne benchmark in Ref. [1] also included the  $F_2^d$  BCDMS deuterium data.

Data Set	Data points	Observable	Ref.
ZEUS97	206	$\tilde{\sigma}^{\text{NC},+}$	[12]
H1lowx97	86	$\tilde{\sigma}^{\text{NC},+}$	[8]
NMC	67	$F_2^p$	[13]
NMC_pd	73	$F_2^d/F_2^p$	[14]
BCDMS	157	$F_2^p$	[15]
Total	589		

Table 3: Data points used in the MSTW benchmark fit after kinematic cuts of  $Q^2 > 9 \text{ GeV}^2$  and  $W^2 > 15 \text{ GeV}^2$  are applied.

- All correlations between systematics are neglected, and statistical and systematic errors are added in quadrature.
- Normalizations of individual data sets are fitted with a rescaling of uncertainties to avoid systematic bias.
- The  $F_2^d/F_2^p$  data are corrected for nuclear shadowing effects [16].

The MSTW analysis presented here makes the same choices as the Thorne benchmark, but with  $\alpha_s(M_Z) = 0.112$ , and additionally

- a global correction of  $-3.4\%$  is applied to the luminosity of the published H1 MB 97 data [8] following a luminosity reanalysis [17].
- a quartic penalty term in the  $\chi^2$  definition is given to normalizations which deviate from the central value.

## 2 Experimental Error Propagation<sup>1</sup>

### 2.1 Introduction

Standard error estimation of proton parton distribution functions (PDFs) relies on the assumption that all errors follow Gaussian (or normal) statistics. However, this assumption may not always be correct. Some systematic uncertainties such as luminosity and detector acceptance follow rather a log-normal distribution (see Section [18]). Compared to the Gaussian case, the lognormal distribution which has the same mean and root mean square (RMS), is asymmetric and has a shifted peak, as shown illustratively in Figure 3. Therefore, the non-Gaussian behaviour of the experimental uncertainties could lead to an additional uncertainty of the resulting PDFs. An alternative to the standard error propagation is a toy Monte Carlo (MC) method. Here, an implementation of the MC method is presented for estimation of the PDF uncertainties with various assumptions for the error distribution. In addition, this MC method provides an independent cross check of the standard error propagation when assuming the Gaussian error distributions.

### 2.2 Method

The Monte Carlo technique consists firstly in preparing replicas of the initial data sets which have the central value of the cross sections,  $\sigma_i$ , fluctuating within its systematic and statistical uncer-

<sup>1</sup>Contributing authors: J. Feltesse, A. Glazov, V. Radescu

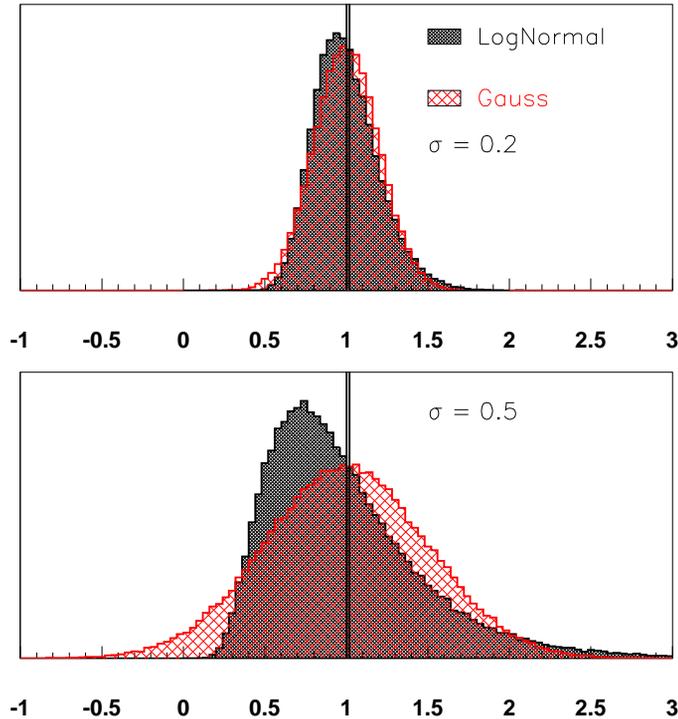


Fig. 3: Comparison of the lognormal (black, darker hatching) and Gaussian (red, lighter hatching) probability distributions. The distributions are shown with mean equal to one, and two different choices for the RMS (for both distribution):  $\sigma = 0.2$  (top) and  $\sigma = 0.5$ .

tainties taking into account all point to point correlations. Various assumptions can be considered for the error distributions. When dealing with the statistical and point to point uncorrelated errors, one could allow each data point to randomly fluctuate within its uncorrelated uncertainty assuming either Gauss, lognormal, or any other desired form of the error distribution. For example, for Gaussian errors

$$\sigma_i \longrightarrow \sigma_i (1 + \delta_i^{uncorr} \cdot R_i), \quad (6)$$

where  $\delta_i^{uncorr}$  corresponds to the uncorrelated uncertainties and  $R_i$  is a random number chosen from a normal distribution with a mean of 0 and a standard deviation of 1. Hence, the central value of each cross section point  $i$  is shifted by  $\delta_i^{uncorr} \cdot R_i$ .

For the systematic errors, the treatment is a bit more complicated than above. This is due to the correlation between data points and that, in general, the data points are sensitive to the systematic sources with a different strength  $\delta_{ij}$ , where index  $i$  ( $j$ ) runs over all the cross section points (all systematic sources). In order to take this into account, for each systematic source  $j$  a uniformly distributed *fluctuation probability*  $P_j$  is selected. Then, for each data point  $i$  the central value of cross section is shifted such that probability of this shift, which depends on  $\delta_{ij}$  and the exact form of the probability distribution function, is equal  $P_j$  (for positive  $\delta_{ij}$ ) or  $(1 - P_j)$  (for negative  $\delta_{ij}$ ). In other words, each central value of the cross section is shifted with

the same probability of the corresponding systematic shift. For example for the Gaussian errors, this procedure is equivalent to

$$\sigma_i \longrightarrow \sigma_i \left( 1 + \delta_i^{uncorr} \cdot R_i + \sum_j^{N_{sys}} \delta_{ij}^{corr} \cdot R_j \right), \quad (7)$$

where in addition to the shifts for the uncorrelated errors previously explained,  $R_j$  corresponds to another random number chosen from a normal distribution with mean of 0 and standard deviation of 1 as a fluctuation for the systematic source  $j$ . Hence, the central values of the cross sections are shifted in addition by  $\delta_{ij}^{corr} \cdot R_j$  for each systematic shift.

This preparation of the data is repeated for  $N$  times, where high statistics is desirable for more accurate results. For this study we used  $N > 100$  which proved to suffice. For each replica, a next to leading order (NLO) QCD fit is performed to extract the PDFs. The errors on the PDFs are estimated from the RMS of the spread of the  $N$  lines corresponding to the  $N$  individual fits to extract PDF.

A fit to the published H1 HERA-I data of neutral and charged current  $e^\pm p$  scattering cross sections [2] using the settings discussed in Sect. 1.1 has been performed, using the QCDNUM program [19].

### 2.3 Validation of the Method

The MC method is tested by comparing the standard error estimation of the PDF uncertainties with the MC techniques by assuming that all the errors (statistical and systematic) follow Gaussian (normal) distribution. Figure 4 shows good agreement between the methods.

### 2.4 Test of various assumptions for the error distributions

Two cases are considered which may represent most likely the error distributions: (1) the log-normal distribution for the luminosity uncertainty and the rest of the errors are set to follow the Gauss shape, (2) the lognormal distributions for all the systematic errors and the statistical errors are set to follow the Gauss distributions. The results for the first case (1) are shown in Figure 5. The results of the tests for the case when lognormal distributions for all the systematic uncertainties are assumed is shown in Figure 5. We observe that for the precise H1 HERA-1 data the effect of using lognormal distribution, which is considered for some systematic uncertainties more physical, is similar to the pure gauss distribution case.

### 2.5 Conclusions

A simple method to estimate PDF uncertainties has been built within QCD Fit framework. Assuming only gauss distribution of all errors, the results agree well with the standard error estimation. This method allows to check the effect of non- gauss assumptions for distributions of the experimental uncertainties. For the H1 data, results are similar to the gauss case when using lognormal. The method could be extended for other physical variables (i.e. cross sections) for cross checks with the standard error evaluation.

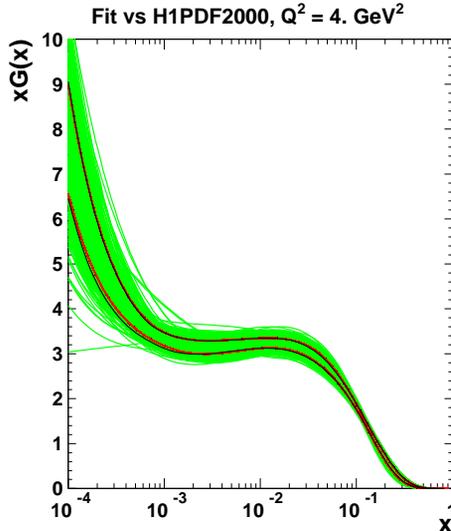


Fig. 4: Comparison between the standard error calculations and the Gauss error distribution is shown for the gluon PDF. Green lines represent the spread of Monte Carlo generated allowances for the errors, and the red lines are the RMS of this spread. The black lines correspond to the standard error calculations of the PDF errors.

### 3 HERA–LHC Benchmark

This benchmark is based on the Alekhin/Thorne benchmark of Ref. [1], whose settings has been given in Sect. 1.2. Both the Alekhin and Thorne fits had the following features:

- uncertainties determined using the Hessian method with  $\Delta\chi^2 = 1$ ;
- input PDFs are parameterized using the following functional form:

$$x f_i(x, Q_0^2) = A_i(1-x)^{b_i}(1 + \epsilon_i x^{0.5} + \gamma_i x)x^{a_i}. \quad (8)$$

with  $\epsilon_i$  and  $\gamma_i$  set to zero for the sea and gluon distributions. Hence, there were a total of 13 free PDF parameters plus  $\alpha_s(M_Z)$  after imposing sum rules.

Here, we reanalyze it within the MSTW and NNPDF approaches. First, we summarize the respective MSTW and NNPDF approaches, and especially their differences when compared to the previous HERALHC benchmark fits of Ref. [1]. Then, results for benchmark fits obtained with the various different approaches are compared to each other. Finally, we compare each benchmark fit to its counterpart based on a wider range of data, i.e. the NNPDF1.0 [5] reference and the MRST01 [20] and MSTW08 [6, 7] PDFs.

#### 3.1 MSTW approach <sup>2</sup>

The benchmark analysis is now much more closely aligned to the global analysis than was the case for the Thorne benchmark compared to the MRST global analysis. It follows the general

<sup>2</sup>Contributing authors: R. S. Thorne, G. Watt

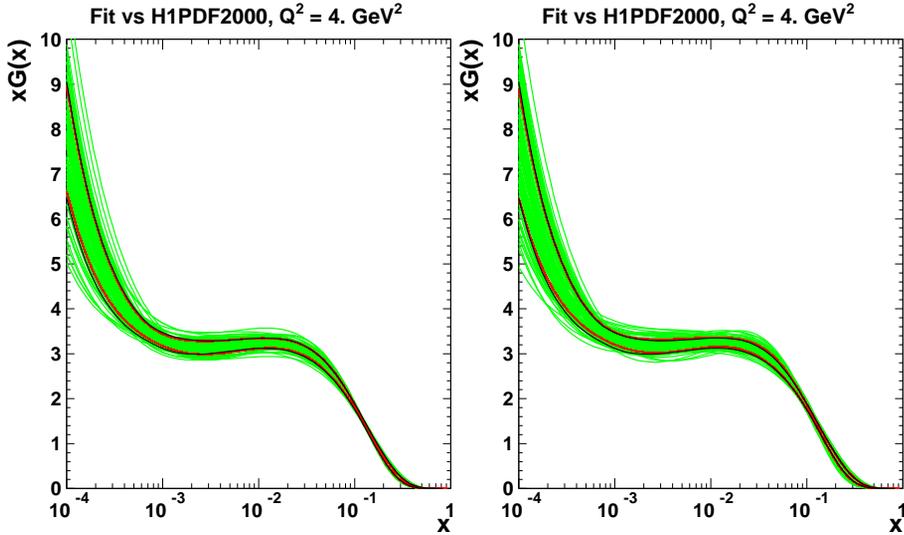


Fig. 5: Comparison between errors on PDFs obtained via standard error calculation (black) where Gauss assumption is used, and errors obtained via Monte Carlo method (red) where luminosity uncertainty is allowed to fluctuate according to lognormal distributions and all the other uncertainties follow the Gaussian distribution (left), and where all the systematic uncertainties are allowed to fluctuate according to lognormal distributions (right). Only the gluon PDF is shown, where the errors are larger. The green lines show the spread of the  $N$  individual fits.

approach taken by the MRST (or more recently, MSTW) group, and is similar to that described in Ref. [20]. There are some new features which are explained below.

- *Input parameterization.* We take the input PDF parameterization at  $Q_0^2 = 1 \text{ GeV}^2$  to be:

$$xu_v(x, Q_0^2) = A_u x^{\eta_1} (1-x)^{\eta_2} (1 + \epsilon_u \sqrt{x} + \gamma_u x), \quad (9)$$

$$xd_v(x, Q_0^2) = A_d x^{\eta_3} (1-x)^{\eta_4} (1 + \epsilon_d \sqrt{x} + \gamma_d x), \quad (10)$$

$$xS(x, Q_0^2) = A_S x^{\delta_S} (1-x)^{\eta_S} (1 + \epsilon_S \sqrt{x} + \gamma_S x), \quad (11)$$

$$xg(x, Q_0^2) = A_g x^{\delta_g} (1-x)^{\eta_g} (1 + \epsilon_g \sqrt{x} + \gamma_g x) + A_{g'} x^{\delta_{g'}} (1-x)^{\eta_{g'}}, \quad (12)$$

where  $S = 2(\bar{u} + \bar{d} + \bar{s})$ ,  $s = \bar{s} = 0.1S$  and  $\bar{d} = \bar{u}$ . The parameters  $A_u$ ,  $A_d$  and  $A_g$  are fixed by sum rules, leaving potentially 19 free parameters. In practice, to reduce the number of highly correlated parameters, making linear error propagation unreliable, we determine the central value of the benchmark fit by freeing all 19 parameters, then fix 6 of those at the best-fit values when calculating the Hessian matrix used to determine the PDF uncertainties, giving a total of 13 eigenvectors. This is the same procedure as used in the MSTW 2008 global fit [6, 7], where there are an additional 3 free parameters associated with  $\bar{d} - \bar{u}$  and an additional 4 free parameters associated with strangeness, giving a total of 20 eigenvectors. Note that the parameterization used in the previous Alekhin/Thorne benchmark fits was considerably more restrictive, where the  $\epsilon_S$ ,  $\gamma_S$ ,  $\epsilon_g$  and  $\gamma_g$  parameters were set to zero, and the second (negative) gluon term was omitted entirely. In addition,  $\epsilon_u$  was held fixed for the Thorne benchmark fit, leaving a total of 12 eigenvectors. We find

that the more flexible gluon parameterization, allowing it to go negative at very small  $x$ , is very highly correlated with the value obtained for  $\alpha_s$ , and a value of  $\alpha_s(M_Z) = 0.105$  is obtained if it is allowed to go free at the same time as the other parameters, therefore we instead choose to fix it at  $\alpha_s(M_Z) = 0.112$  as in the NNPDF benchmark fit.

- *Error propagation.* Apart from the more flexible input parameterization, the other major difference in the new MSTW version of the HERA–LHC benchmark fit, with respect to the previous Thorne (MRST) version, is the choice of tolerance,  $T = \sqrt{\Delta\chi^2}$ . The MRST benchmark fit used the standard choice  $T = 1$  for one-sigma uncertainties. More precisely, the distance  $t$  along each normalized eigenvector direction was taken to be 1, and ideal quadratic behaviour about the minimum was assumed, giving  $T \approx t = 1$ . The MRST global fit used  $T = \sqrt{50}$  for a 90% confidence level (C.L.) uncertainty band; however, this is not appropriate when fitting a smaller number of data sets. Recently, a new procedure has been developed [6, 7] which enables a *dynamic* determination of the tolerance for each eigenvector direction, by demanding that each data set must be described within its one-sigma (or 90%) C.L. limits according to a hypothesis-testing criterion, after rescaling the  $\chi^2$  for each data set so that the value at the global minimum corresponds to the most probable value. Application of this procedure to the MSTW benchmark fit gives  $T \sim 3$  for one-sigma uncertainties and  $T \sim 5$  for 90% C.L. uncertainties. For the MSTW global fit, the typical values of  $T$  required are slightly larger, with more variation between different eigenvector directions. The increase in  $T$  in the global fit is mainly due to the inclusion of some less compatible data sets, while the greater variation in  $T$  between eigenvectors is due to the fact that some parameters, particularly those associated with  $s$  and  $\bar{s}$ , are constrained by far fewer data sets than others. In the MSTW fits, the data set normalizations are allowed to vary, with the aforementioned penalty term, when determining the PDF uncertainties. For global fits this automatically leads to a small increase in uncertainty compared to the MRST determinations, where data set normalisations were held fixed when calculating the Hessian matrix used for error propagation. In the MRST benchmark fit the data set normalizations were allowed to vary. To calculate the uncertainty bands from the eigenvector PDF sets, we use the formula for asymmetric errors given, for example, in Eq. (13) of Ref. [20].

### 3.2 NNPDF approach<sup>3</sup>

The NNPDF approach was proposed in Ref. [21], and it was applied there and in Ref. [22] to the parameterization of the structure function  $F_2(x, Q^2)$  with only two or more experimental data sets respectively. In Ref. [23] it was first used for the determination of a single PDF (the isotriplet quark distribution), and in Ref. [5] a full set of PDFs fit based on DIS data (NNPDF1.0) was presented. Because the method has been discussed extensively in these references, here we only summarize briefly its main features.

- *Error propagation.* We make a Monte Carlo sample of the probability distribution of the experimental data by generating an ensemble of  $N$  replicas of artificial data following a

---

<sup>3</sup>Contributing authors: R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione, J. Rojo, M. Ubiali

multi-gaussian distribution centered on each data point with full inclusion of the experimental covariance matrix. Each replica is used to construct a set of PDFs, thereby propagating the statistical properties of the data Monte Carlo sample to a final Monte Carlo sample of PDFs. Here we shall take  $N = 100$ . The method is the same as discussed in Sect. 2.2, the only difference being the treatment of normalization errors: relative normalizations are fitted in the H1 approach, while they are included among the systematic errors in the Monte Carlo data generation in the NNPDF approach (see Refs. [2, 5] for details of the respective procedures) .

- *Input parameterization.* Each PDF is parameterized with a functional form provided by a neural network. The architecture for the neural network is the same for all PDFs, and yields a parameterization with 37 free parameters for each PDF. This is a very redundant parameterization, it is chosen in order to avoid parameterization bias; neural networks are a particularly convenient way of dealing with redundant parameterizations. Note that sum rules are also imposed.
- *Minimization.* A redundant parameterization allows for fitting not only the underlying physical behaviour, but also statistical noise. Therefore, the minimization is stopped not at the absolute minimum of the  $\chi^2$ , but rather before one starts fitting noise. This optimal stopping point is determined as follows: the data in each replica are randomly assigned either to a training or to a validation set. The fit is performed on data of the training set only, while the validation set is used as a monitor. The fit is stopped when the quality of the fit to the training set keeps improving, but the quality of the fit to the validation set deteriorates.

### 3.3 Comparison between the Benchmark Parton Distributions

Data Set	$\chi_{\text{bench}}^2/N_{\text{data}}$	$\chi_{\text{global}}^2/N_{\text{data}}$
ZEUS97	1.09	1.18
H1lowx97	1.03	1.00
NMC	1.40	1.45
NMC <sub>pd</sub>	1.24	1.32
BCDMS	1.21	1.98
Total	1.19	1.53

Table 4: NNPDF  $\chi^2$  for the total and each single data set, both for the benchmark and global fit.

The  $\chi^2$  per data point for the NNPDF and MSTW fits are shown in Table 4 and 5 respectively. Note that in the MSTW fit statistical and systematic errors are added in quadrature, so the quantity shown is the diagonal contribution to the  $\chi^2$ . The quality of the NNPDF is seen to be uniformly good. The quality of the MSTW is also uniform, though it cannot be compared directly because of the different way systematics are treated. The comparison of each benchmark fit to the corresponding global fit will be discussed in Sect. 3.4 below.

In Fig. 6 the PDFs from the NNPDF and MSTW benchmark fits presented here are compared to those by Thorne from Ref. [1] at the same reference scale of  $Q^2 = 20 \text{ GeV}^2$  used there

Data set	$\chi_{\text{bench}}^{\text{diag } 2} / N_{\text{data}}$	$\chi_{\text{global}}^{\text{diag } 2} / N_{\text{data}}$
ZEUS97	0.76	0.79
H1lowx97	0.53	0.54
NMC	1.08	1.11
NMC_pd	0.78	0.89
BCDMS	0.74	1.13
Total	0.76	0.89

Table 5: MSTW  $\chi^2$  for the total and each single data set, both for the benchmark and global fit. Notice that statistical and systematic errors are added in quadrature and that relative data set normalizations are fitted.

(denoted as MRST01 in the figure). The benchmark fit by Alekhin [1] is not shown as the PDFs are very close to the those by Thorne displayed in Fig. 6.

For PDFs and kinematical regions where data are available, namely the small- $x$  gluon and sea quark and the large- $x$   $u_v$  distributions, the central values of the NNPDF fit are quite close to those of the MRST and MSTW fits, despite the differences in methodology. The central values of the PDFs are slightly different for the MRST and MSTW benchmark fits due to the use of BCDMS  $F_2^d$  data in the former, which affects mainly valence quarks. Where extrapolation is needed, such as for the  $d_v$  distribution, which is constrained only by the small amount of data on the ratio  $F_2^d/F_2^p$ , or the large- $x$  sea quark, central values are rather more different (though the Alekhin/MRST/MSTW benchmark central values are within the NNPDF error band). The exception is the smallest- $x$  gluon, where the form of the MSTW parameterization results in a very sharp turn-over. However, even here the uncertainty bands are close to overlapping.

Differences are sizeable in the estimation of uncertainties. Firstly, uncertainty bands for NNPDF benchmark are significantly larger than for the MSTW benchmark, which in turn are in general somewhat larger than those for the MRST benchmark. The difference between MRST and MSTW, which are based on similar methodology, is due to use of a dynamic tolerance and a more flexible gluon parameterization in MSTW (see Sect. 3.1). Secondly, the width of the uncertainty band for NNPDF benchmark varies rather more than that of the MRST benchmark according to the PDF and the kinematic region, though this is not quite so much the case comparing to MSTW benchmark. Indeed, the NNPDF uncertainties are quite small in the region between  $x = 0.01$  and  $x = 0.1$  (where there is the bulk of HERA and fixed-target data), while they blow up in the large- $x$  region for the sea quark or the small- $x$  gluon, where there is less or no experimental information. The smallness of the uncertainty band for MSTW for the small- $x$  valence quarks may be partially due to the lack of flexibility in the parameterization: note that because of sum rules, the size of uncertainties in the data and extrapolation region are correlated.

Finally, the MRST/MSTW central value generally falls within the NNPDF uncertainty band, but the NNPDF central value tends to fall outside the MRST/MSTW uncertainty band whenever the central values differ significantly.

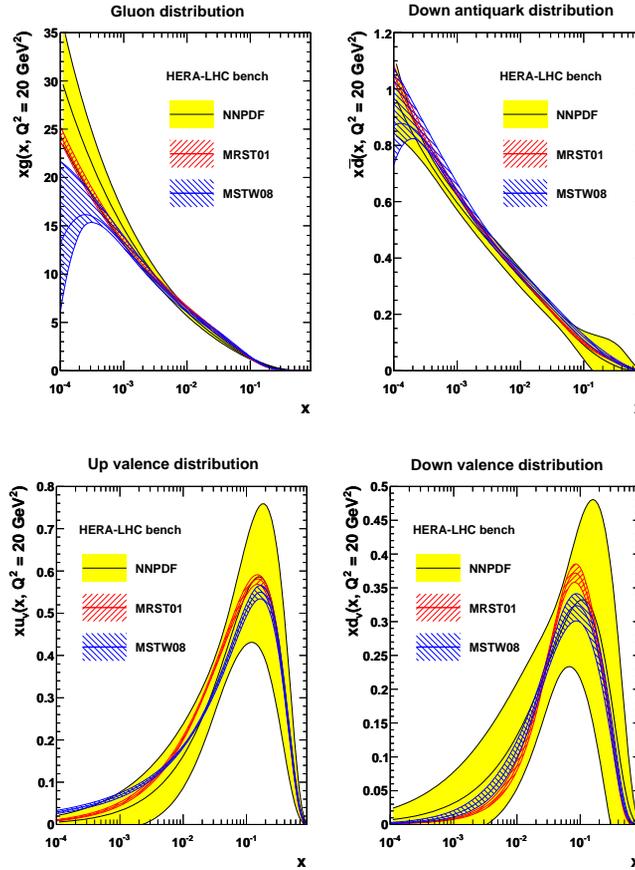


Fig. 6: Comparison of the NNPDF, MRST and MSTW benchmark fits for the gluon,  $d$ -sea,  $u$ -valence and  $d$ -valence at  $Q^2 = 20 \text{ GeV}^2$ . All uncertainties shown correspond to one- $\sigma$  bands.

### 3.4 Comparison of the Benchmark Parton Distributions and Global Fits

In Fig. 7 we compare the NNPDF benchmark fit to the NNPDF1.0 reference fit of Ref. [5] (NNPDF global, henceforth), while in Fig. 8 we compare the MSTW benchmark fit to the MRST01 [20] (MRST global, henceforth) and MSTW08 [6, 7] global fits (MSTW global, henceforth).

The  $\chi^2$  of the NNPDF benchmark and global fits are compared in Table 4, while those of the MSTW benchmark and global fits are compared in Table 5. Note that for the NNPDF fits the  $\chi^2$  is computed using the full covariance matrix, while for the MSTW fits systematic and statistical uncertainties are added in quadrature. Note also that the MRST and MSTW global fits are carried out in a general-mass variable flavour number scheme rather than the zero-mass variable flavour number scheme used in the corresponding benchmark fits, whereas for NNPDF both global and benchmark fits are done with a zero-mass variable flavour number scheme. Comparison of the quality of each benchmark to the corresponding global fit to the same points in Table 5 shows a significant deterioration in the quality of the fit (total  $\Delta\chi^2 \gg 1$ ), especially for

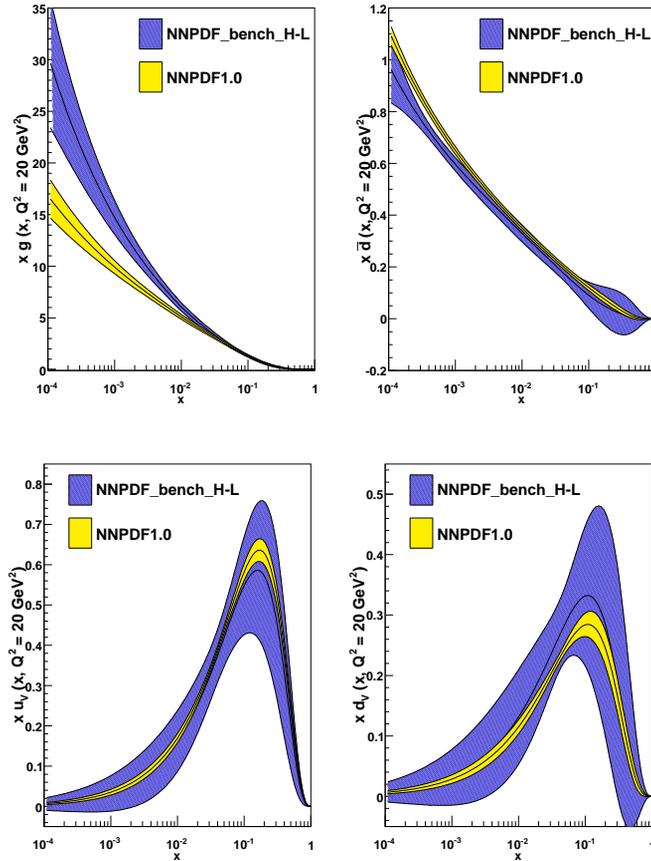


Fig. 7: Comparison of the NNPDF benchmark and reference fits for the gluon,  $d$ -sea,  $u$ -valence and  $d$ -valence at  $Q^2 = 20 \text{ GeV}^2$ .

the BCDMS  $F_2^p$  data. All fits appear to be acceptable for all data sets: for instance, even though the  $\chi^2$  of the NNPDF global fit for the benchmark subset of data is 1.98, it is equal to 1.59 [5] for the full BCDMS set of data. However, the increase in  $\chi^2$  suggests that there might be data inconsistencies.

Let us now compare each pair of benchmark and global fits. For NNPDF, the difference in central value between benchmark and reference is comparable to that found between the MRST or Alekhin global fits and their benchmark counterparts in Ref. [1]. However, the NNPDF global and benchmark fits remain compatible within their respective error bands. Indeed, the NNPDF benchmark fit has a rather larger error band than the reference, as one would expect from a fit based on a rather smaller set of (compatible) data. Such a behaviour was however not observed in the comparison between global and benchmark MRST and Alekhin fits of Ref. [1].

It is interesting to observe that the gluon shape at low  $x$  of the benchmark and global NNPDF disagree at the one  $\sigma$  level (though they agree at two  $\sigma$ ). This can be understood as a consequence of the fact that the value of  $\alpha_s$  in the two fits is sizably different ( $\alpha_s = 0.112$  vs.

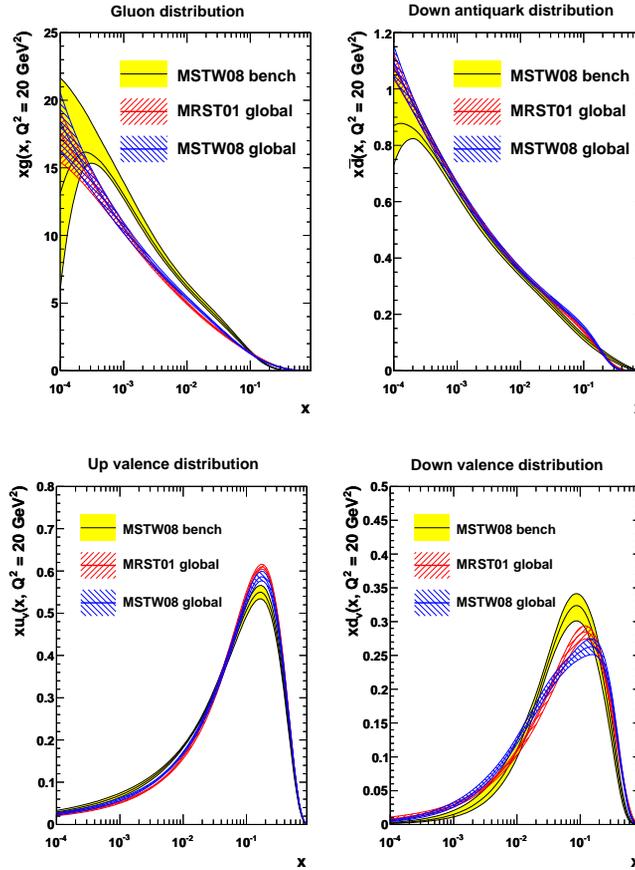


Fig. 8: Comparison of the MSTW benchmark and MRST/MSTW global fits for the gluon,  $d$ -sea,  $u$ -valence and  $d$ -valence at  $Q^2 = 20 \text{ GeV}^2$ . All uncertainties shown correspond to one- $\sigma$  bands.

$\alpha_s = 0.119$ ). Theoretical uncertainties related to the value of  $\alpha_s$  were shown in Ref. [5] to be negligible and thus not included in the NNPDF error band, but of course they become relevant if  $\alpha_s$  is varied by several standard deviations ( $3.5 \sigma$ , in this case).

Coming now to MSTW, we first notice that, as discussed in Sect. 3.3, the MSTW benchmark set has somewhat larger uncertainty bands than the MRST benchmark set and thus also than each of the sets obtained from global fits. Consequently, the MSTW benchmark PDFs are generally far more consistent with the MSTW global fit sets than the corresponding comparison between MRST benchmark PDFs and global fit PDFs shown in Ref. [1], largely due to the more realistic uncertainties in the MSTW benchmark. Comparing central values we see exactly the same feature in the gluon distribution as the NNPDF group, and the explanation is likewise the same, highlighting possible difficulties in comparing PDFs obtained with different values of  $\alpha_s(M_Z)$ .

Unlike for the NNPDF group, the MSTW group sees some degree of incompatibility between the benchmark PDFs and the global fit PDFs for the valence quarks, particularly in the case

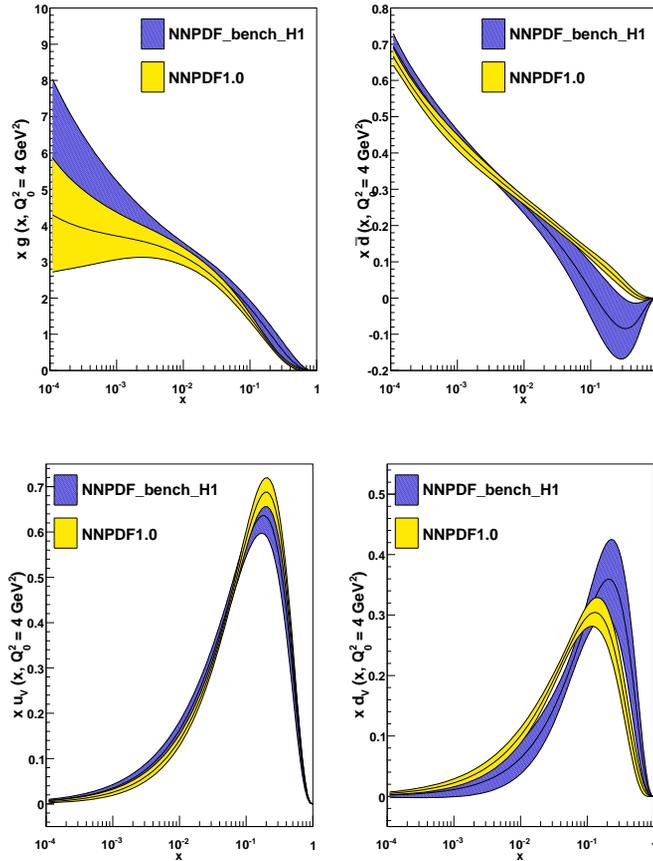


Fig. 9: Comparison of the NNPDF benchmark and reference fits for the gluon,  $d$ -sea,  $u_v$  and  $d_v$  at  $Q^2 = 4 \text{ GeV}^2$ .

of the down valence. This may be related to the assumption  $\bar{u} = \bar{d}$ , which constrains valence quarks and sea quarks in an artificial manner since there is less flexibility to alter each independently. Indeed, in the global fits there is an excess of  $\bar{d}$  over  $\bar{u}$  which maximizes at  $x = 0.1$ . Forcing equivalence of antiquark distributions might therefore lead to a deficit of down sea quarks and a corresponding excess of up sea quarks, and also, for the same reason, to an excess of down valence quarks. These are indeed seen both in the NNPDF and MSTW benchmark fits when compared to the respective global fits. The effect is however well within the uncertainty bands for NNPDF, which indeed do not observe any statistically significant difference between results of a fit to the reduced benchmark data set with the  $\bar{u} = \bar{d}$  assumption (as presented in Fig. 7) or without it (as presented in Ref. [5], Fig. 12).

As well as this important effect one sees that the main discrepancy at  $x = 0.1$  for down valence quarks is greater when comparing the benchmark fits to the global MSTW fit than to the global MRST fit. This is because recent new Tevatron data on  $Z$  rapidity distributions and lepton asymmetry from  $W$  decays provide a strong constraint on the down quark, and some of this new data shows considerable tension with other data sets.

## 4 H1 Benchmark

We now discuss the extension of the fit using the settings of Sect. 1.1 to also include the NNPDF approach. Results are compared both to those of the NNPDF reference fit, and to those obtained by the H1 fit of Sect. 2 to the same data. We then compare the NNPDF benchmark and reference, with the specific aim of addressing the issue of the dependence of the results on the size of the data set (H1 dataset vs. the HERA–LHC dataset of Sect. 3). Finally, the H1 and NNPDF benchmark fits are compared to each other with the purpose of understanding the impact of the respective methodologies.

### 4.1 NNPDF analysis<sup>4</sup>

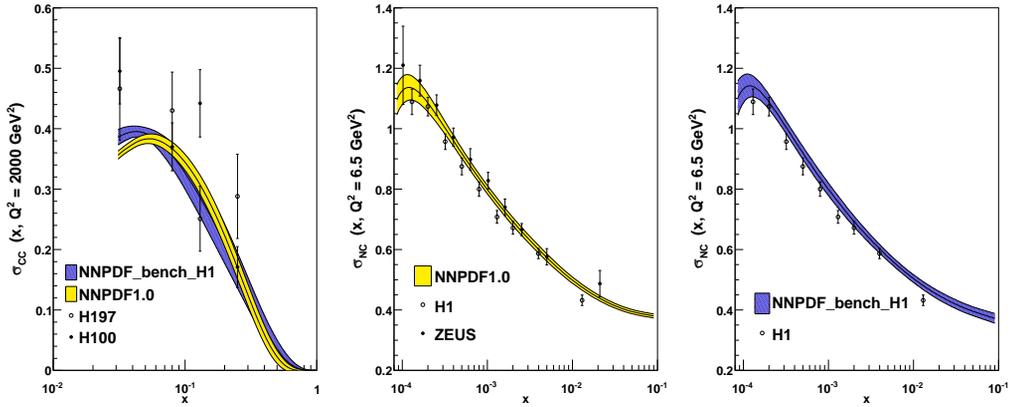


Fig. 10: Left: NNPDF benchmark and reference fits at  $\sqrt{s} = 301 \text{ GeV}$  compared to H1 charged current data. Center: NNPDF reference fit compared to H1 and ZEUS neutral current data. Right: NNPDF benchmark fit compared to H1 neutral current data.

The results of the NNPDF benchmark are compared to the NNPDF reference fit results in Fig. 9. The general features of the benchmark are analogous to those of the HERA–LHC benchmark discussed in Section 3.4, with some effects being more pronounced because the benchmark dataset is now even smaller. Specifically, we observe that uncertainties bands blow up when data are removed: this is very clear for instance in the  $\bar{d}$  distribution at large- $x$ , as a consequence of the fact that the benchmark dataset of Table 1 does not include deuterium data. The negative value of this PDF at large  $x$  is presumably unphysical and it would disappear if positivity of charged current cross sections were imposed, including also the (anti-)neutrino ones. The only positivity constraint in the NNPDF fit is imposed on the  $F_L$  structure function [5], because this is the only DIS observable whose positivity is not constrained by the full data set.

It is interesting to note however that this effect is not observed for the  $u_v$  distribution, where instead the benchmark and the reference fit show almost equal uncertainties. In order to

<sup>4</sup>Contributing authors R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, A. Piccione, J. Rojo, M. Ubiali

understand this, in Fig. 10 we compare two situations with or without error shrinking, by examining the predictions obtained using the benchmark and reference fits for some observables to the corresponding data. A first plot (left) shows the shrinking of the uncertainty on the prediction for the charged-current cross section in the reference fit. This is mostly due to the CHORUS neutrino data, which are in the reference and not in the benchmark. These data are clearly consistent with the H1 data shown in the plot. The subsequent pair of plots compares (center) the prediction for the neutral-current cross section from the reference fit compared to H1 and ZEUS data (both of which are used for the reference fit), and (right) from the benchmark fit to the H1 data only (which are the only ones used in the benchmark fit). The uncertainty bands in the two fits are similar size: indeed, the ZEUS and H1 data display a systematic disagreement which is approximately the size of this uncertainty band. Hence, the (small but significant) systematic inconsistency between the ZEUS and H1 data prevents reduction of the uncertainty band when the ZEUS data are added to the fit, beyond the size of this discrepancy. Therefore, the NNPDF methodology leads to combined uncertainties for inconsistent data which are similar to those obtained with the so-called PDG (or scale-factor) method [24].

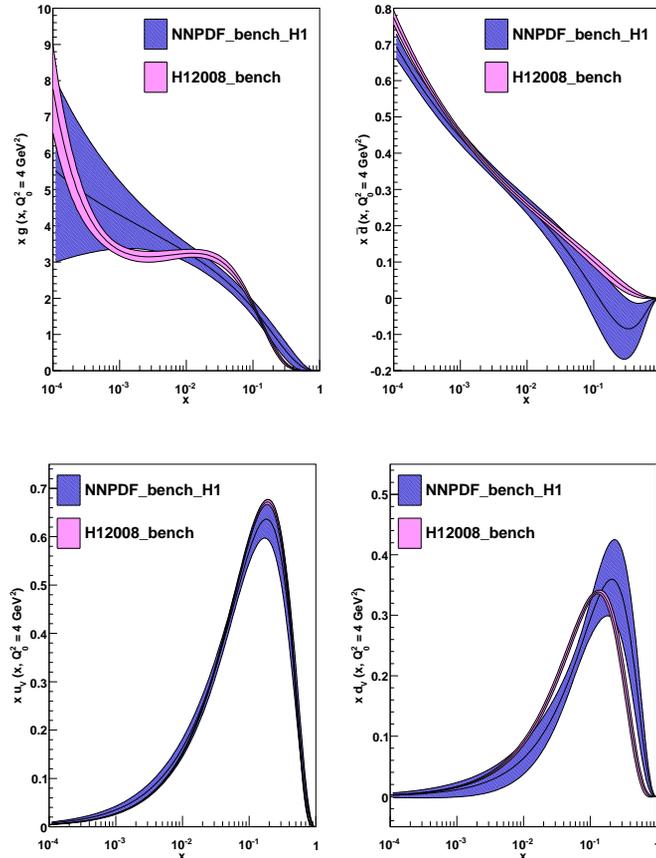


Fig. 11: Comparison of the NNPDF and H1 benchmark fit for the gluon,  $d$ -sea,  $u_v$  and  $d_v$  at  $Q^2 = 4 \text{ GeV}^2$ .

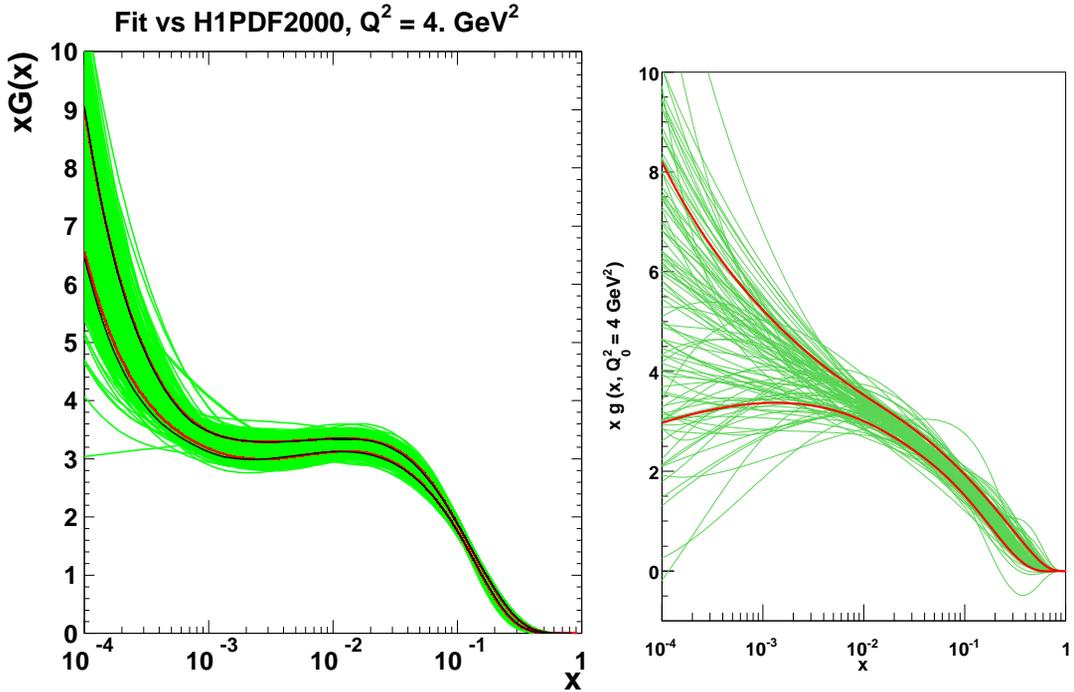


Fig. 12: The Monte Carlo set of gluon PDFs for the H1 benchmark (left, same as Fig. 4) and the NNPDF benchmark. The red lines show the one-sigma contour calculated from the Monte Carlo set, and in the H1 case the black lines show the Hessian one-sigma contour.

Data Set	$\chi^2_{\text{H1}}/N_{\text{data}}$	$\chi^2_{\text{NNPDF}}/N_{\text{data}}$
H197mb	0.83	0.82
H197lowQ2	0.90	0.87
H197NC	0.69	0.80
H197CC	0.73	0.97
H199NC	0.88	1.01
H199CC	0.62	0.84
H199NChy	0.35	0.35
H100NC	0.97	1.00
H100CC	1.07	1.38
Total	0.88	0.96

Table 6: H1 and NNPDF  $\chi^2$  for the total and each single data set. Cross correlations among data sets are neglected to evaluate the  $\chi^2$  of a single data set.

Notice that if relative normalization are fitted (as done by in the H1 approach of Sect. 2) instead of being treated simply as a source of systematics, this systematic inconsistency would be significantly reduced in the best-fit. The associate uncertainty however then appears as an

addition source of systematics. This happens when H1 and ZEUS data are combined in a single dataset (see Section [18] below). In the NNPDF approach, instead, this systematics is produced by the Monte Carlo procedure.

## 4.2 Comparison between the Benchmark Parton Distributions

The  $\chi^2$  of the H1 and NNPDF benchmarks are given in Table 6, while the corresponding PDFs are compared in Fig. 11. Furthermore, in Fig. 12 we show the respective full Monte Carlo PDF sets in the case of the gluon distribution.

The quality of the two fits is comparable, the differences in  $\chi^2$  being compatible with statistical fluctuations. In the region where experimental information is mostly concentrated, specifically for the  $u_v$  distribution over all the  $x$ -range and for the  $\bar{d}$  and the  $d_v$  distributions in the small- $x$  range, the results of the two fits are in good agreement, though the H1 uncertainty bands are generally somewhat smaller.

In the region where experimental information is scarce or missing, sizable differences are found, similar to those observed when comparing the MRST/MSTW bench and NNPDF bench to the HERA–LHC benchmark of Sect. 3.3. Specifically, in these regions NNPDF uncertainties are generally larger than H1 bands: the width of the uncertainty band for the H1 fit varies much less between the data and extrapolation regions than that of the NNPDF bench. Also, the H1 central value always falls within the NNPDF uncertainty band, but the NNPDF central value tends to fall outside the H1 uncertainty band whenever the central values differ significantly. Figure 12 suggests that this may be due to the greater flexibility of the functional form in the NNPDF fit. Specifically, the  $\bar{d}$  quark distribution at large  $x$  does not become negative in the H1 fit, because this behaviour is not allowed by the parameterization.

## References

- [1] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.
- [2] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C30**, 1 (2003). hep-ex/0304003.
- [3] W. T. Giele and S. Keller, *Phys. Rev.* **D58**, 094023 (1998). hep-ph/9803393.
- [4] W. T. Giele, S. A. Keller, and D. A. Kosower (2001). hep-ph/0104052.
- [5] NNPDF Collaboration, R. D. Ball *et al.*, *Nucl. Phys.* **B809**, 1 (2009). 0808.1231.
- [6] G. Watt, A. D. Martin, W. J. Stirling, and R. S. Thorne, *Recent Progress in Global PDF Analysis*. Preprint 0806.4890, 2008.
- [7] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, *Parton Distributions for the LHC*. Preprint 0706.0456, 2007.
- [8] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C21**, 33 (2001). hep-ex/0012053.
- [9] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C13**, 609 (2000). hep-ex/9908059.

- [10] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C19**, 269 (2001). hep-ex/0012052.
- [11] W. Giele *et al.*, *The QCD/SM working group: Summary report*. Preprint hep-ph/0204316, 2002.
- [12] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C21**, 443 (2001). hep-ex/0105090.
- [13] New Muon Collaboration Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B483**, 3 (1997). hep-ph/9610231.
- [14] New Muon Collaboration Collaboration, M. Arneodo *et al.*, Nucl. Phys. **B487**, 3 (1997). hep-ex/9611022.
- [15] BCDMS Collaboration, A. C. Benvenuti *et al.*, Phys. Lett. **B223**, 485 (1989).
- [16] B. Badelek and J. Kwiecinski, Phys. Rev. **D50**, 4 (1994). hep-ph/9401314.
- [17] A. d. R. Vargas Trevino, *Measurement of the inclusive  $e p$  scattering cross section at low  $Q^2$  and  $x$  at HERA*. Prepared for 15th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS2007), Munich, Germany, 16-20 Apr 2007.
- [18] *Extraction of the proton pdfs from a combined fit of h1 and zeus inclusive dis cross sections*. These proceedings.
- [19] M. Botje, *Qcdnum16.12*. Available from <http://www.nikhef.nl/h24/qcdnum/>.
- [20] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C28**, 455 (2003). hep-ph/0211080.
- [21] S. Forte, L. Garrido, J. I. Latorre, and A. Piccione, JHEP **05**, 062 (2002). hep-ph/0204232.
- [22] NNPDF Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, JHEP **03**, 080 (2005). hep-ph/0501067.
- [23] NNPDF Collaboration, L. Del Debbio, S. Forte, J. I. Latorre, A. Piccione, and J. Rojo, JHEP **03**, 039 (2007). hep-ph/0701127.
- [24] Particle Data Group Collaboration, C. Amsler *et al.*, Phys. Lett. **B667**, 1 (2008).

# Determination of parton distributions

A. Cooper-Sarkar, A. Glazov, G. Li, J. Grebenyuk, V. Lendermann

## 1 Extraction of the proton PDFs from a combined fit of H1 and ZEUS inclusive DIS cross sections <sup>1</sup>

### 1.1 Introduction

The kinematics of lepton hadron scattering is described in terms of the variables  $Q^2$ , the invariant mass of the exchanged vector boson, Bjorken  $x$ , the fraction of the momentum of the incoming nucleon taken by the struck quark (in the quark-parton model), and  $y$  which measures the energy transfer between the lepton and hadron systems. The differential cross-section for the neutral current (NC) process is given in terms of the structure functions by

$$\frac{d^2\sigma(e^\pm p)}{dx dQ^2} = \frac{2\pi\alpha^2}{Q^4 x} [Y_+ F_2(x, Q^2) - y^2 F_L(x, Q^2) \mp Y_- xF_3(x, Q^2)],$$

where  $Y_\pm = 1 \pm (1 - y)^2$ . The structure functions  $F_2$  and  $xF_3$  are directly related to quark distributions, and their  $Q^2$  dependence, or scaling violation, is predicted by perturbative QCD. For low  $x$ ,  $x \leq 10^{-2}$ ,  $F_2$  is sea quark dominated, but its  $Q^2$  evolution is controlled by the gluon contribution, such that HERA data provide crucial information on low- $x$  sea-quark and gluon distributions. At high  $Q^2$ , the structure function  $xF_3$  becomes increasingly important, and gives information on valence quark distributions. The charged current (CC) interactions also enable us to separate the flavour of the valence distributions at high- $x$ , since their (LO) cross-sections are given by,

$$\begin{aligned} \frac{d^2\sigma(e^+p)}{dx dQ^2} &= \frac{G_F^2 M_W^4}{(Q^2 + M_W^2)^2 2\pi x} x [(\bar{u} + \bar{c}) + (1 - y)^2(d + s)], \\ \frac{d^2\sigma(e^-p)}{dx dQ^2} &= \frac{G_F^2 M_W^4}{(Q^2 + M_W^2)^2 2\pi x} x [(u + c) + (1 - y)^2(\bar{d} + \bar{s})]. \end{aligned}$$

Parton Density Function (PDF) determinations are usually obtained in global NLO QCD fits [1–3], which use fixed target DIS data as well as HERA data. In such analyses, the high statistics HERA NC  $e^+p$  data have determined the low- $x$  sea and gluon distributions, whereas the fixed target data have determined the valence distributions. Now that high- $Q^2$  HERA data on NC and CC  $e^+p$  and  $e^-p$  inclusive double differential cross-sections are available, PDF fits can be made to HERA data alone, since the HERA high  $Q^2$  cross-section data can be used to determine the valence distributions. This has the advantage that it eliminates the need for heavy target corrections, which must be applied to the  $\nu$ -Fe and  $\mu D$  fixed target data. Furthermore there is no need to assume isospin symmetry, i.e. that  $d$  in the proton is the same as  $u$  in the neutron, since the  $d$  distribution can be obtained directly from CC  $e^+p$  data.

The H1 and ZEUS collaborations have both used their data to make PDF fits [3], [4]. Both of these data sets have very small statistical uncertainties, so that the contribution of systematic

---

<sup>1</sup>Contributing authors: A. Cooper-Sarkar, A. Glazov, G. Li for the H1-ZEUS combination group.

uncertainties becomes dominant and consideration of point to point correlations between systematic uncertainties is essential. The ZEUS analysis takes account of correlated experimental systematic errors by the Offset Method, whereas H1 uses the Hessian method [5]. Whereas the resulting ZEUS and H1 PDFs are compatible, the gluon PDFs have rather different shapes, see Fig 7, and the uncertainty bands spanned by these analyses are comparable to those of the global fits.

It is possible to improve on this situation since ZEUS and H1 are measuring the same physics in the same kinematic region. These data have been combined using a ‘theory-free’ Hessian fit in which the only assumption is that there is a true value of the cross-section, for each process, at each  $x, Q^2$  point [6]. Thus each experiment has been calibrated to the other. This works well because the sources of systematic uncertainty in each experiment are rather different, such that all the systematic uncertainties are re-evaluated. The resulting correlated systematic uncertainties on each of the combined data points are significantly smaller than the statistical errors. This combined data set has been used as the input to an NLO QCD PDF fit. The consistency of the input data set and its small systematic uncertainties enables us to calculate the experimental uncertainties on the PDFs using the  $\chi^2$  tolerance,  $\Delta\chi^2 = 1$ . This represents a further advantage compared to the global fit analyses where increased tolerances of  $\Delta\chi^2 = 50 - 100$  are used to account for data inconsistencies.

For the HERAPDF0.1 fit presented here, the role of correlated systematic uncertainties is no longer crucial since these uncertainties are relatively small. This ensures that similar results are obtained using either Offset or Hessian methods, or by simply combining statistical and systematic uncertainties in quadrature. The  $\chi^2$  per degree of freedom for a Hessian fit is 553/562 and for a quadrature fit it is 428/562. For our central fit we have chosen to combine the 43 systematic uncertainties which result from the separate ZEUS and H1 data sets in quadrature, and to Offset the 4 sources of uncertainty which result from the combination procedure. The  $\chi^2$  per degree of freedom for this fit is 477/562. This procedure results in the most conservative estimates on the resulting PDFs as illustrated in Fig. 1 which compares the PDFs and their experimental uncertainties as evaluated by the procedure of our central fit and as evaluated by treating the 47 systematic uncertainties by the Hessian method.

Despite this conservative procedure, the experimental uncertainties on the resulting PDFs are impressively small and a thorough consideration of further uncertainties due to model assumptions is necessary. In Section 1.2 we briefly describe the data combination procedure. In Section 1.3 we describe the NLO QCD analysis and model assumptions. In Section 1.4 we give results. In Section 1.5 we give a summary of the fit results and specifications for release of the HERAPDF0.1 to LHAPDF. In Section 1.6 we investigate the predictions of the HERAPDF0.1 for  $W$  and  $Z$  cross-sections at the LHC.

## 1.2 Data Combination

The data combination is based on assumption that the H1 and ZEUS experiments measure the same cross section at the same kinematic points. The systematic uncertainties of the measurements are separated, following the prescription given by the H1 and ZEUS, into point to point correlated sources  $\alpha_j$  and uncorrelated systematic uncertainty, which is added to the statistical uncertainty in quadrature to result in total uncorrelated uncertainty  $\sigma_i$  for each bin  $i$ . The

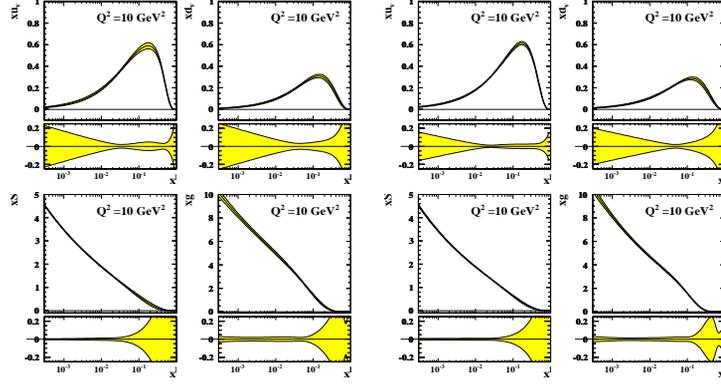


Fig. 1: HERAPDFs,  $xu_v$ ,  $xd_v$ ,  $xS$ ,  $xg$  at  $Q^2 = 10\text{GeV}^2$ . (Left) with experimental uncertainties evaluated as for the central fit (see text) and (right) with experimental uncertainties evaluated by accounting for the 47 systematic errors by the Hessian method.

correlated systematic sources are considered to be uncorrelated between H1 and ZEUS. All uncertainties are treated as multiplicative i.e. proportional to the central values, which is a good approximation for the measurement of the cross sections.

A correlated probability distribution function for the physical cross sections  $M^{i,\text{true}}$  and systematic uncertainties  $\alpha_{j,\text{true}}$  for a single experiment corresponds to a  $\chi^2$  function:

$$\chi_{\text{exp}}^2(M^{i,\text{true}}, \alpha_{j,\text{true}}) = \sum_i \frac{\left[ M^{i,\text{true}} - \left( M^i + \sum_j \frac{\partial M^i}{\partial \alpha_j} \frac{M^{i,\text{true}}}{M^i} (\alpha_{j,\text{true}}) \right) \right]^2}{\left( \sigma_i \frac{M^{i,\text{true}}}{M^i} \right)^2} + \sum_j \frac{(\alpha_{j,\text{true}})^2}{\sigma_{\alpha_j}^2}, \quad (1)$$

where  $M^i$  are the central values measured by the experiment,  $\partial M^i / \partial \alpha_j$  are the sensitivities to the correlated systematic uncertainties and  $\sigma_{\alpha_j}$  are the uncertainties of the systematic sources. For more than one experiment, total  $\chi_{\text{tot}}^2$  can be represented as a sum of  $\chi_{\text{exp}}^2$ . The combination procedure allows to represent  $\chi_{\text{tot}}^2$  in the following form:

$$\chi_{\text{tot}}^2(M^{i,\text{true}}, \beta_{j,\text{true}}) = \chi_0^2 + \sum_i \frac{\left[ M^{i,\text{true}} - \left( M^{i,\text{ave}} + \sum_j \frac{\partial M^{i,\text{ave}}}{\partial \beta_j} \frac{M^{i,\text{true}}}{M^{i,\text{ave}}} (\beta_{j,\text{true}}) \right) \right]^2}{\left( \sigma_{i,\text{ave}} \frac{M^{i,\text{true}}}{M^{i,\text{ave}}} \right)^2} + \sum_j \frac{(\beta_{j,\text{true}})^2}{\sigma_{\beta_j}^2}. \quad (2)$$

Here the sum runs over a union set of the cross section bins. The value of the  $\chi_{\text{tot}}^2$  at the minimum,  $\chi_0^2$ , quantifies consistency of the experiments.  $M^{i,\text{ave}}$  are the average values of the cross sections and  $\beta_j$  correspond to the new systematic sources which can be obtained from the original sources  $\alpha_j$  through the action of an orthogonal matrix. In essence, the average of several data sets allows

one to represent the total  $\chi^2$  in a form which is similar to that corresponding to a single data set, Eq. 1, but with modified systematic sources.

The combination is applied to NC and CC cross section data taken with  $e^+$  and  $e^-$  beams simultaneously to take into account correlation of the systematic uncertainties. The data taken with proton beam energies of  $E_p = 820$  GeV and  $E_p = 920$  GeV are combined together for inelasticity  $y < 0.35$ , for this a small center of mass energy correction is applied. For the combined data set there are 596 data points and 43 experimental systematic sources. The  $\chi_0^2/dof = 510/599$  is below 1, which indicates conservative estimation of the uncorrelated systematics.

Besides the experimental uncertainties, four additional sources related to the assumptions made for the systematic uncertainties are considered. Two of the extra sources deal with correlation of the H1 and ZEUS data for estimation of the photoproduction background and simulation of hadronic energy scale. These sources introduce additional  $\sim 1\%$  uncertainty for  $y > 0.6$  and  $y < 0.02$  data. The third source covers uncertainty arising from the center of mass correction by varying  $F_L = F_L^{QCD}$  to  $F_L = 0$ . The resulting uncertainty reaches few per mille level for  $y \sim 0.35$ . Finally, some of the systematic uncertainties, for example background subtraction, may not be necessary multiplicative but rather additive, independent of the cross section central values. The effect of additive assumption for the errors is evaluated by comparing the average obtained using Eq. 1 and an average in which  $M^{i,true}/M^{i,ave}$  scaling is removed for all but global normalization errors.

### 1.3 QCD Analysis

The QCD predictions for the structure functions are obtained by solving the DGLAP evolution equations [7–9] at NLO in the  $\overline{MS}$  scheme with the renormalisation and factorization scales chosen to be  $Q^2$ <sup>2</sup>. The DGLAP equations yield the PDFs at all values of  $Q^2$  provided they are input as functions of  $x$  at some input scale  $Q_0^2$ . This scale has been chosen to be  $Q_0^2 = 4\text{GeV}^2$  and variation of this choice is considered as one of the model uncertainties. The resulting PDFs are then convoluted with NLO coefficient functions to give the structure functions which enter into the expressions for the cross-sections. The choice of the heavy quark masses is,  $m_c = 1.4$ ,  $m_b = 4.75\text{GeV}$ , and variation of these choices is included in the model uncertainties. For this preliminary analysis, the heavy quark coefficient functions have been calculated in the zero-mass variable flavour number scheme. The strong coupling constant was fixed to  $\alpha_s(M_Z^2) = 0.1176$  [12], and variations in this value of  $\pm 0.002$  have also been considered.

The fit is made at leading twist. The HERA data have a minimum invariant mass of the hadronic system,  $W^2$ , of  $W_{min}^2 = 300\text{ GeV}^2$  and a maximum  $x$ ,  $x_{max} = 0.65$ , such that they are in a kinematic region where there is no sensitivity to target mass and large- $x$  higher twist contributions. However a minimum  $Q^2$  cut is imposed to remain in the kinematic region where perturbative QCD should be applicable. This has been chosen to be  $Q_{min}^2 = 3.5\text{ GeV}^2$ . Variation of this cut is included as one of the model uncertainties.

A further model uncertainty is the choice of the initial parameterization at  $Q_0^2$ . Three types of parameterization have been considered. For each of these choices the PDFs are parameterized

<sup>2</sup>The programme QCDNUM [10] has been used and checked against the programme QCDfit [11].

by the generic form

$$xf(x) = Ax^B(1-x)^C(1 + Dx + Ex^2 + Fx^3), \quad (3)$$

and the number of parameters is chosen by 'saturation of the  $\chi^2$ ', such that parameters  $D, E, F$  are only varied if this brings significant improvement to the  $\chi^2$ . Otherwise they are set to zero.

The first parameterization considered follows that used by the ZEUS collaboration. The PDFs for  $u$  valence,  $xu_v(x)$ ,  $d$  valence,  $xd_v(x)$ , total sea,  $xS(x)$ , the gluon,  $xg(x)$ , and the difference between the  $d$  and  $u$  contributions to the sea,  $x\Delta(x) = x(\bar{d} - \bar{u})$ , are parameterized.

$$\begin{aligned} xu_v(x) &= A_{uv}x^{B_{uv}}(1-x)^{C_{uv}}(1 + D_{uv}x + E_{uv}x^2) \\ xd_v(x) &= A_{dv}x^{B_{dv}}(1-x)^{C_{dv}} \\ xS(x) &= A_Sx^{B_S}(1-x)^{C_S} \\ xg(x) &= A_gx^{B_g}(1-x)^{C_g}(1 + D_gx) \\ x\Delta(x) &= A_\Delta x^{B_\Delta}(1-x)^{C_\Delta} \end{aligned}$$

The total sea is given by,  $xS = 2x(\bar{u} + \bar{d} + \bar{s} + \bar{c} + \bar{b})$ , where  $\bar{q} = q_{sea}$  for each flavour,  $u = u_v + u_{sea}$ ,  $d = d_v + d_{sea}$  and  $q = q_{sea}$  for all other flavours. There is no information on the shape of the  $x\Delta$  distribution in a fit to HERA data alone and so this distribution has its parameters fixed, such that its shape is consistent with Drell-Yan data and its normalization is consistent with the size of the Gottfried sum-rule violation. A suppression of the strange sea with respect to the non-strange sea of a factor of 2 at  $Q_0^2$ , is imposed consistent with neutrino induced dimuon data from NuTeV. The normalisation parameters,  $A_{uv}, A_{dv}, A_g$ , are constrained to impose the number sum-rules and momentum sum-rule. The  $B$  parameters,  $B_{uv}$  and  $B_{dv}$  are set equal, since there is no information to constrain any difference. Finally this ZEUS-style parameterization has eleven free parameters.

The second parameterization considered follows that of the H1 Collaboration. The choice of quark PDFs which are parameterized is different. The quarks are considered as  $u$ -type and  $d$ -type,  $xU = x(u_v + u_{sea} + c)$ ,  $xD = x(d_v + d_{sea} + s)$ ,  $x\bar{U} = x(\bar{u} + \bar{c})$  and  $x\bar{D} = x(\bar{d} + \bar{s})$ , assuming  $q_{sea} = \bar{q}$ , as usual. These four (anti-)quark distributions are parameterized separately.

$$\begin{aligned} xU(x) &= A_Ux^{B_U}(1-x)^{C_U}(1 + D_Ux + E_Ux^2 + F_Ux^3) \\ xD(x) &= A_Dx^{B_D}(1-x)^{C_D}(1 + D_Dx) \\ x\bar{U}(x) &= A_{\bar{U}}x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}} \\ x\bar{D}(x) &= A_{\bar{D}}x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}} \\ xg(x) &= A_gx^{B_g}(1-x)^{C_g} \end{aligned}$$

Since the valence distributions must vanish as  $x \rightarrow 0$ , the parameters,  $A$  and  $B$  are set equal for  $xU$  and  $x\bar{U}$ ;  $A_U = A_{\bar{U}}$ ,  $B_U = B_{\bar{U}}$ ; and for  $xD$  and  $x\bar{D}$ ;  $A_D = A_{\bar{D}}$ ,  $B_D = B_{\bar{D}}$ . Since there is no information on the flavour structure of the sea it is also necessary to set  $B_{\bar{U}} = B_{\bar{D}}$ , such that there is a single  $B$  parameter for all four quark distributions. The normalisation,  $A_g$ , of the gluon

is determined from the momentum sum-rule and the parameters  $D_U$  and  $D_D$  are determined by the number sum-rules. Assuming that the strange and charm quark distributions can be expressed as  $x$  independent fractions,  $f_s = 0.33$  and  $f_c = 0.15$ , of the  $d$  and  $u$  type sea respectively, gives the further constraint  $A_{\bar{U}} = A_{\bar{D}}(1 - f_s)/(1 - f_c)$ , which ensures that  $\bar{u} = \bar{d}$  at low  $x$ . Finally this H1-style parameterization has 10 free parameters.

The third parameterization we have considered combines the best features of the previous two. It has less model dependence than the ZEUS-style parameterization in that it makes fewer assumptions on the form of sea quark asymmetry  $x\Delta$ , and it has less model dependence than the H1-style parameterization in that it does not assume equality of all  $B$  parameters. Furthermore, although all types of parameterization give acceptable  $\chi^2$  values, the third parameterization has the best  $\chi^2$  and it gives the most conservative experimental errors. This is the parameterization which we chose for our central fit. The PDFs which are parameterized are  $xu_v$ ,  $xd_v$ ,  $xg$  and  $x\bar{U}$ ,  $x\bar{D}$ .

$$\begin{aligned} xu_v(x) &= A_{uv}x^{B_{uv}}(1-x)^{C_{uv}}(1 + D_{uv}x + E_{uv}x^2) \\ xd_v(x) &= A_{dv}x^{B_{dv}}(1-x)^{C_{dv}} \\ x\bar{U}(x) &= A_{\bar{U}}x^{B_{\bar{U}}}(1-x)^{C_{\bar{U}}} \\ x\bar{D}(x) &= A_{\bar{D}}x^{B_{\bar{D}}}(1-x)^{C_{\bar{D}}} \\ xg(x) &= A_gx^{B_g}(1-x)^{C_g} \end{aligned}$$

The normalisation parameters,  $A_{uv}$ ,  $A_{dv}$ ,  $A_g$ , are constrained to impose the number sum-rules and momentum sum-rule. The  $B$  parameters,  $B_{uv}$  and  $B_{dv}$  are set equal,  $B_{uv} = B_{dv}$  and the  $B$  parameters  $B_{\bar{U}}$  and  $B_{\bar{D}}$  are also set equal,  $B_{\bar{U}} = B_{\bar{D}}$ , such that there is a single  $B$  parameter for the valence and another different single  $B$  parameter for the sea distributions. Assuming that the strange and charm quark distributions can be expressed as  $x$  independent fractions,  $f_s = 0.33$  and  $f_c = 0.15$ , of the  $d$  and  $u$  type sea, gives the further constraint  $A_{\bar{U}} = A_{\bar{D}}(1 - f_s)/(1 - f_c)$ . The value of  $f_s = 0.33$  has been chosen to be consistent with determinations of this fraction using neutrino induced di-muon production. This value has been varied to evaluate model uncertainties. The charm fraction has been set to be consistent with dynamic generation of charm from the start point of  $Q^2 = m_c^2$ , in a zero-mass-variable-flavour-number scheme. A small variation of the value of  $f_c$  is included in the model uncertainties. Finally this parameterization has 11 free parameters.

It is well known that the choice of parameterization can affect both PDF shapes and the size of the PDF uncertainties. Fig 2 compares the PDFs and their uncertainties as evaluated using these three different parameterizations. As mentioned earlier, the third parameterization results in the most conservative uncertainties.

We present results for the HERA PDFs based on the third type of parameterization, including six sources of model uncertainty as specified in Table 1. We also compare to results obtained by varying  $\alpha_s(M_Z^2)$  and by varying the choice of parameterization to those of the ZEUS and the H1 styles of parameterization.

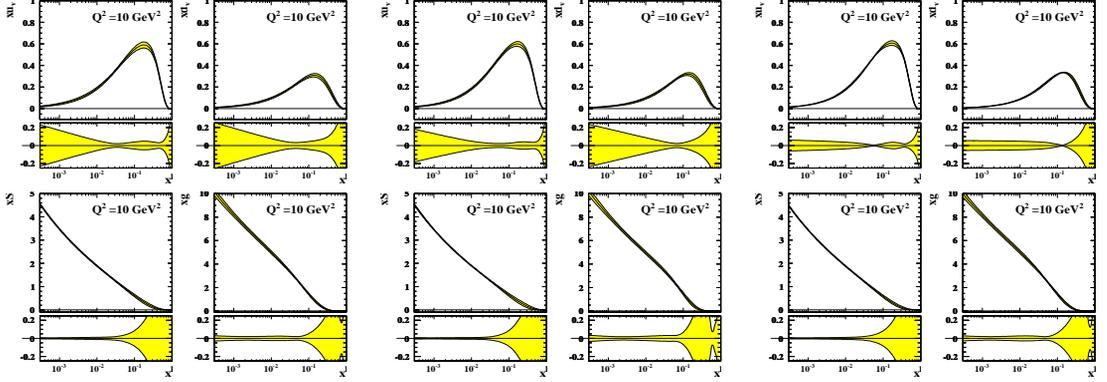


Fig. 2: HERAPDFs,  $xu_v$ ,  $xd_v$ ,  $xS$ ,  $xg$  and their uncertainties at  $Q^2 = 10\text{GeV}^2$ . (Left) for the central fit; (centre) for the ZEUS-style parameterization; (right) for the H1-style parameterization

Model variation	Standard value	Upper Limit	Lower limit
$m_c$	1.4	1.35	1.5
$m_b$	4.75	4.3	5.0
$Q_{min}^2$	3.5	2.5	5.0
$Q_0^2$	4.0	2.0	6.0
$f_s$	0.33	0.25	0.40
$f_c$	0.15	0.12	0.18

Table 1: Standard values of input parameters and cuts, and the variations considered to evaluate model uncertainty

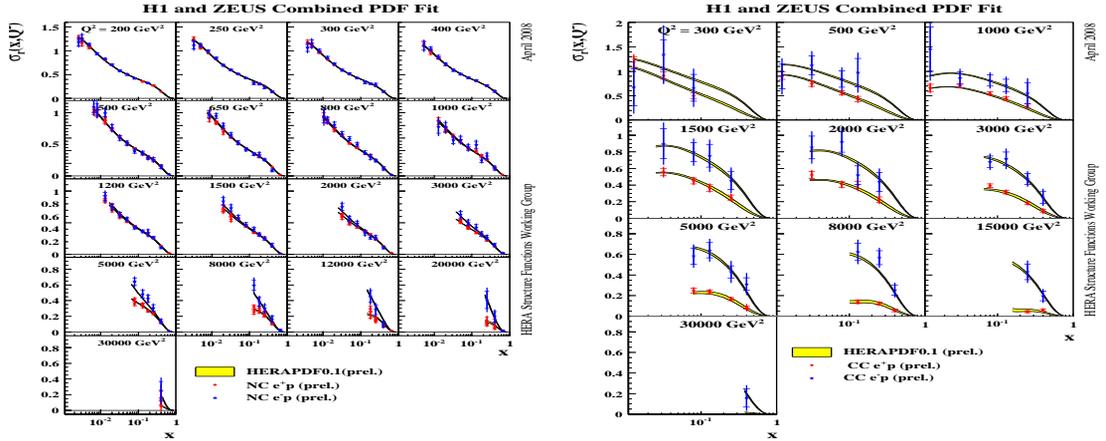


Fig. 3: HERA combined NC (left) and CC (right) data. The predictions of the HERAPDF0.1 fit are superimposed. The uncertainty bands illustrated derive from both experimental and model sources

## 1.4 Results

In Fig. 3 we show the HERAPDF0.1 superimposed on the combined data set for NC data and CC data. In Fig 4 we show the NC data at low  $Q^2$ , and we illustrate scaling violation by showing the reduced cross-section vs.  $Q^2$  for a few representative  $x$  bins. The predictions of the HERAPDF0.1 fit are superimposed, together with the predictions of the ZEUS-JETS and H1PDF2000 PDFs.

Fig. 5 shows the HERAPDF0.1 PDFs,  $xu_v, xd_v, xS, xg$ , as a function of  $x$  at the starting scale  $Q^2 = 4 \text{ GeV}^2$  and at  $Q^2 = 10 \text{ GeV}^2$ . Fig. 6 shows the same PDFs at the scales  $Q^2 = 100, 10000 \text{ GeV}^2$ . Fractional uncertainty bands are shown beneath each PDF. The experimental and model uncertainties are shown separately. As the PDFs evolve with  $Q^2$  the total uncertainty becomes impressively small.

The total uncertainty of the PDFs obtained from the HERA combined data set is much reduced compared to the PDFs extracted from the analyses of the separate H1 and ZEUS data sets, as can be seen from the summary plot Fig. 7, where these new HERAPDF0.1 PDFs are compared to the ZEUS-JETS and H1PDF2000 PDFs. It is also interesting to compare the present HERAPDF0.1 analysis of the combined HERA-I data set with an analysis of the separate data sets which uses the same parameterization and assumptions. Fig 8 makes this comparison. It is clear that it is the data combination, and not the choice of parameterization and assumptions, which has resulted in reduced uncertainties for the low- $x$  gluon and sea PDFs.

The break-up of the HERAPDFs into different flavours is illustrated in Fig. 9, where the PDFs  $xU, xD, x\bar{U}, x\bar{D}$  and  $x\bar{u}, x\bar{d}, x\bar{c}, x\bar{s}$  are shown at  $Q^2 = 10 \text{ GeV}^2$ . The model uncertainty on these PDFs from variation of  $Q_{min}^2, Q_0^2, m_c$  and  $m_b$  is modest. The model uncertainty from variation of  $f_s$  and  $f_c$  is also modest except for its obvious effect on the charm and strange quark distributions.

It is also interesting to look at the results obtained from using the ZEUS-style and H1

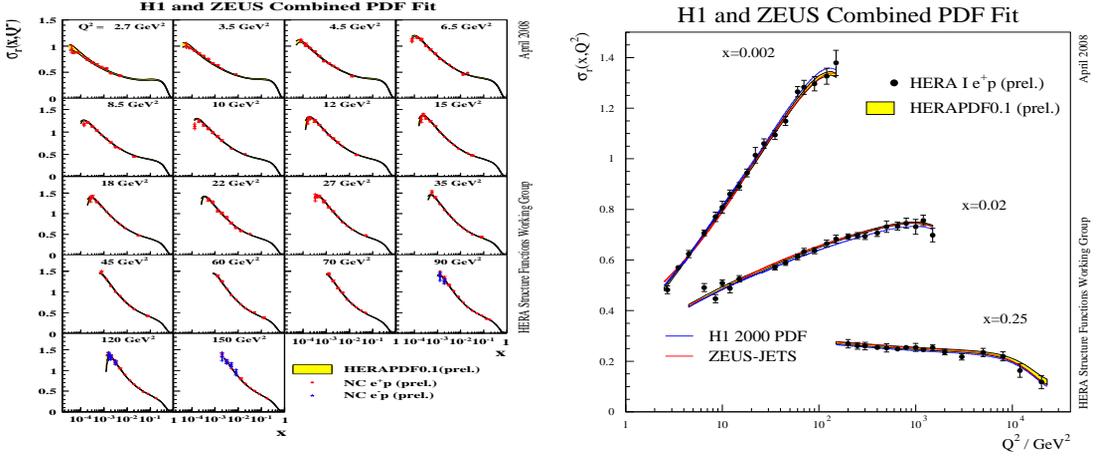


Fig. 4: Left: HERA combined NC data at low  $Q^2$ . Right: the NC reduced cross-section vs  $Q^2$  for three  $x$ -bins. The predictions of the HERAPDF0.1 fit are superimposed, together with the predictions of the ZEUS-JETS and H1PDF2000 PDFs

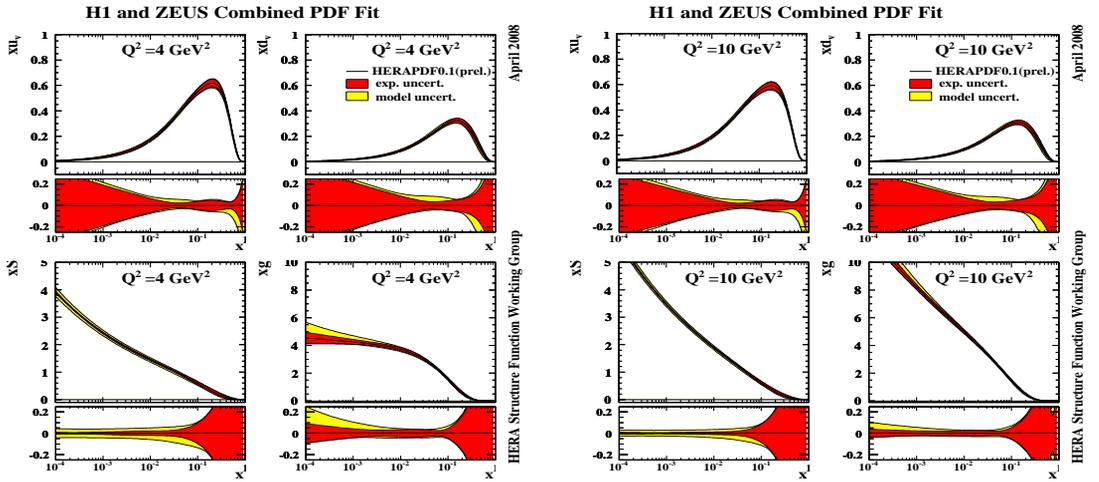


Fig. 5: HERAPDFs,  $xu_v$ ,  $xd_v$ ,  $xS$ ,  $xg$ , at (left)  $Q^2 = 4 \text{ GeV}^2$  and (right)  $Q^2 = 10 \text{ GeV}^2$ . Fractional uncertainty bands are shown beneath each PDF. The experimental and model uncertainties are shown separately as the red and yellow bands respectively

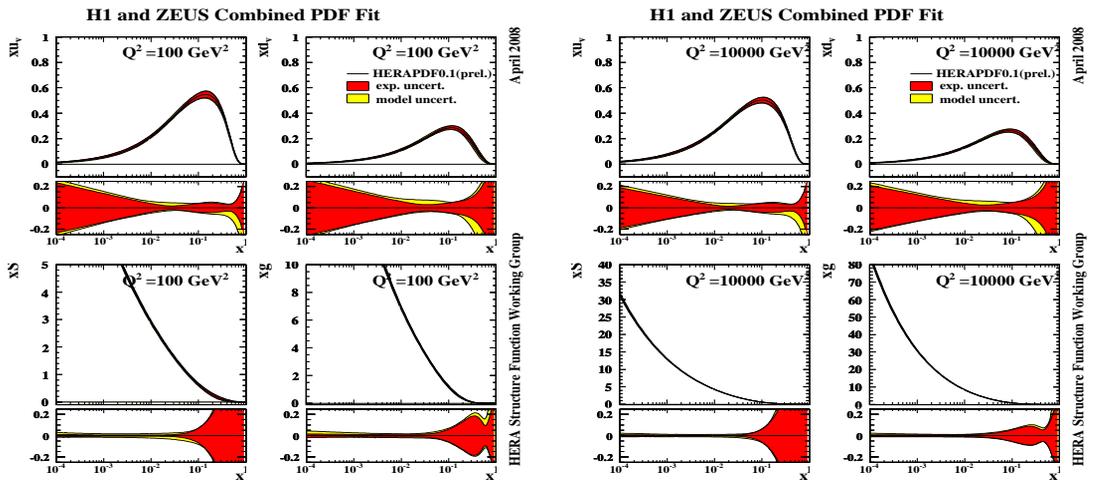


Fig. 6: HERAPDFs,  $xu_v$ ,  $xd_v$ ,  $xS$ ,  $xg$ , at (left)  $Q^2 = 100 \text{ GeV}^2$  and (right)  $Q^2 = 10000 \text{ GeV}^2$ . Fractional uncertainty bands are shown beneath each PDF. The experimental and model uncertainties are shown separately as the red and yellow bands respectively

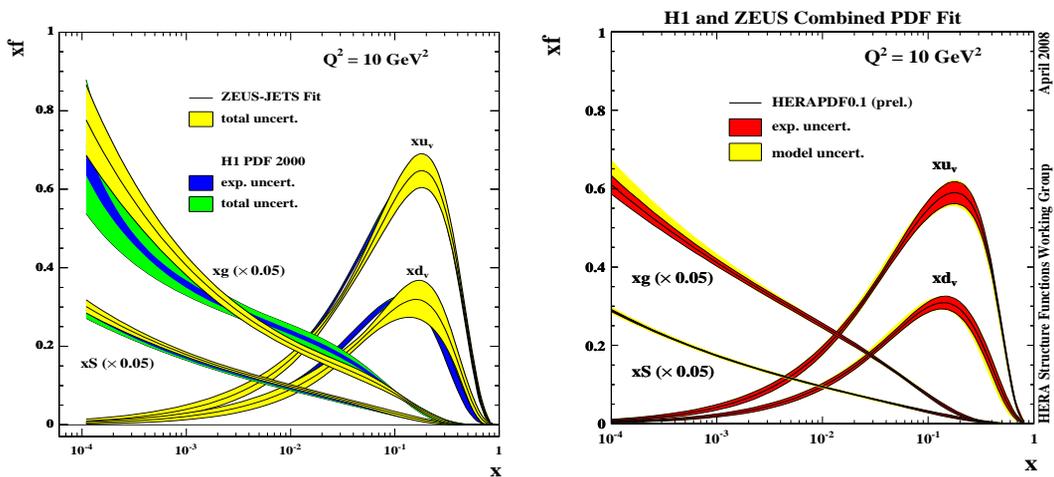


Fig. 7: Left: PDFs from the ZEUS-JETS and H1PDF2000 PDF separate analyses of ZEUS and H1. Right: HERAPDF0.1 PDFs from the analysis of the combined data set

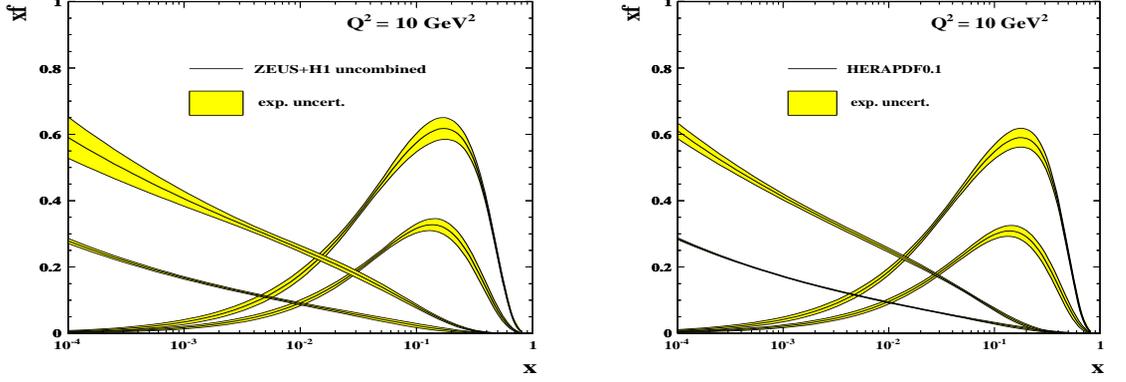


Fig. 8: Left: PDFs resulting from an analysis of the H1 and ZEUS separate data sets using the same parameterization and assumptions as HERAPDF0.1. Right: HERAPDF0.1 PDFs from the analysis of the combined data set (experimental uncertainties only)

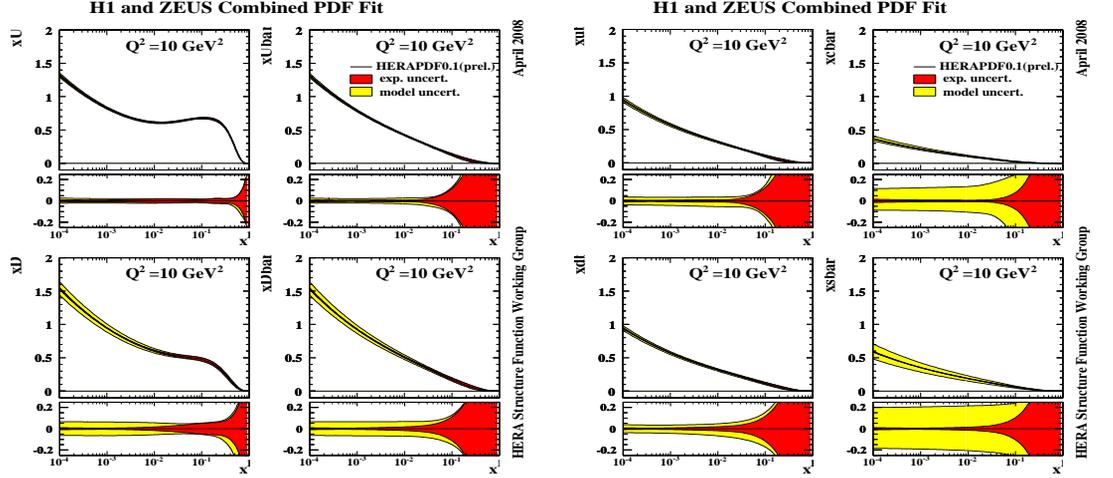


Fig. 9: HERAPDFs at  $Q^2 = 10\text{GeV}^2$ : (left)  $xU, xD, x\bar{U}, x\bar{D}$ ; (right)  $x\bar{u}, x\bar{d}, x\bar{c}, x\bar{s}$ . Fractional uncertainty bands are shown beneath each PDF. The experimental and model uncertainties are shown separately as the red and yellow bands respectively

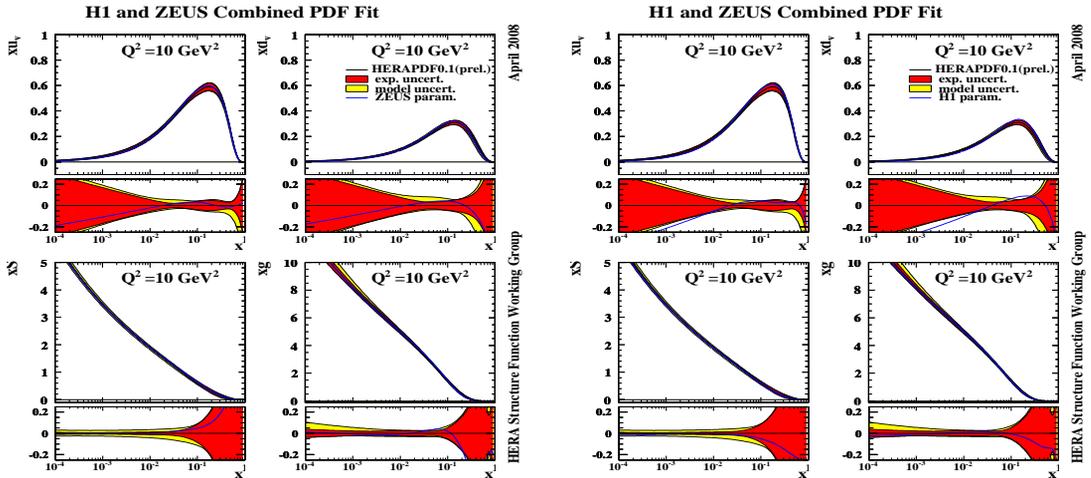


Fig. 10: HERAPDFs at  $Q^2 = 10\text{GeV}^2$ : with the results for the ZEUS-style parameterization (left) and for the H1-style parameterization (right) superimposed as a blue line.

style parameterizations described in Section 1.3. In Fig. 10 these alternative parameterizations are shown as a blue line superimposed on the HERAPDF0.1 PDFs. These variations in parameterization produce changes in the resulting PDFs which are comparable to the experimental uncertainties in the measured kinematic range. A further variation of parameterization originates from the fact that, if the  $D$  parameter for the gluon is allowed to be non-zero, then each type of parameterization yields a double minimum in  $\chi^2$  such that the gluon may take a smooth or a 'humpy' shape. Although the lower  $\chi^2$  is obtained for the smooth shape, the  $\chi^2$  for the 'humpy' shape is still acceptable. The PDFs for the 'humpy' version of our chosen form of parameterization are compared to the standard version in Fig. 11, where they are shown as a blue line superimposed on the HERAPDF0.1 PDFs. This comparison is shown at  $Q^2 = 4\text{GeV}^2$ , where the difference is the greatest. Nevertheless the resulting PDFs are comparable to those of the standard choice. This explains a long-standing disagreement in the shape of the gluon obtained by the separate ZEUS-JETS and H1PDF200 analyses. The ZEUS data favoured the smooth shape and the H1 data favoured the 'humpy' shape. However the precision of the combined data set results in PDFs for these shapes which are not significantly different in the measured kinematic region.

It is also interesting to compare the PDFs for the standard choice to those obtained with a different input value of  $\alpha_s(M_Z^2)$ . The uncertainty on the current PDG value of  $\alpha_s(M_Z^2)$  is  $\pm 0.002$  and thus we vary our central choice by this amount. The results are shown in Fig. 12, where we can see that this variation only affects the gluon PDF, such that the larger(smaller) value of  $\alpha_s(M_Z^2)$  results in a harder(softer) gluon as predicted by the DGLAP equations. The change is outside total uncertainty bands of the standard fit. Finally, Figs. 13 and 14 compare the HERAPDF0.1 PDFs to those of the CTEQ and the MRST/MSTW groups respectively. The uncertainty bands of the CTEQ and MRST/MSTW analyses have been scaled to represent 68% CL limits for direct comparability to the HERAPDF0.1. The HERAPDF0.1 analysis has much

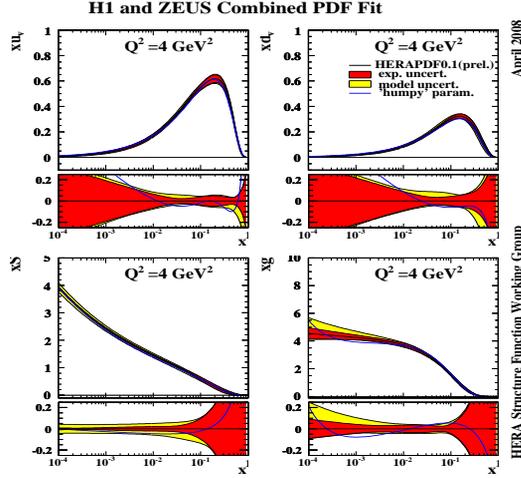


Fig. 11: HERAPDFs at  $Q^2 = 4\text{GeV}^2$ : with the results for the humpy version superimposed as a blue line.

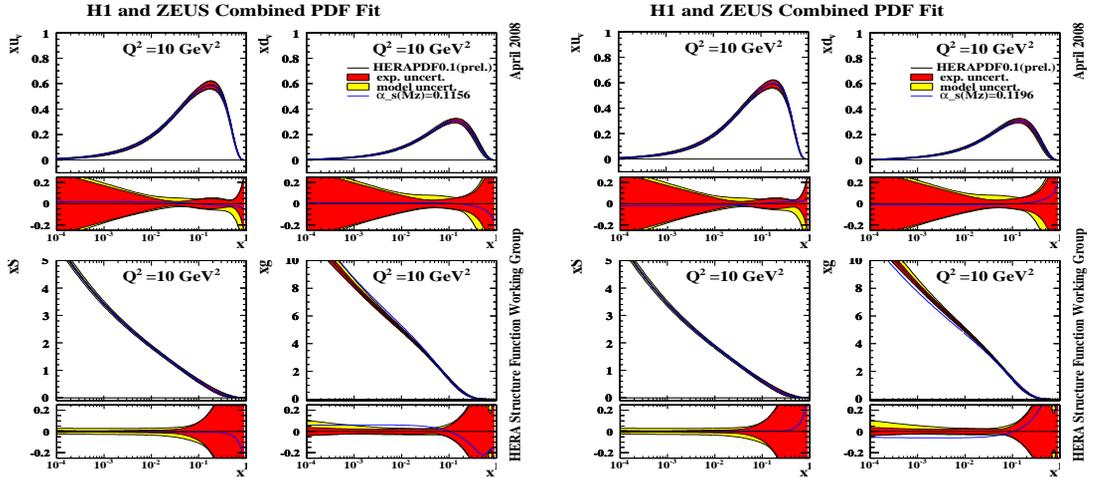


Fig. 12: HERAPDFs at  $Q^2 = 10\text{GeV}^2$ : with the results for  $\alpha_s(M_Z^2) = 0.1156$  (left) and for  $\alpha_s(M_Z^2) = 0.1196$  (right) superimposed as a blue line.

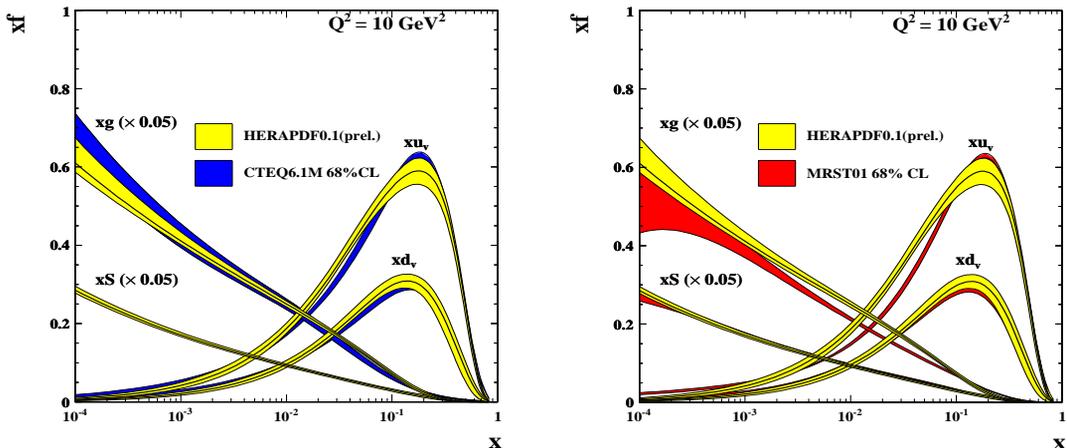


Fig. 13: HERAPDFs at  $Q^2 = 10\text{GeV}^2$  compared to the PDFs from CTEQ6.1 and MRST01

improved precision on the low- $x$  gluon.

### 1.5 Summary of HERAPDF0.1 results

Now that high- $Q^2$  HERA data on NC and CC  $e^+p$  and  $e^-p$  inclusive double differential cross-sections are available, PDF fits can be made to HERA data alone, since the HERA high  $Q^2$  cross-section data can be used to determine the valence distributions and HERA low  $Q^2$  cross-section data can be used to determine the Sea and gluon distributions. The combined HERA-I data set, of neutral and charged current inclusive cross-sections for  $e^+p$  and  $e^-p$  scattering, has been used as the sole input for an NLO QCD PDF fit in the DGLAP formalism. The consistent treatment of systematic uncertainties in the joint data set ensures that experimental uncertainties on the PDFs can be calculated without need for an increased  $\chi^2$  tolerance. This results in PDFs with greatly reduced experimental uncertainties compared to the separate analyses of the ZEUS and H1 experiments. Model uncertainties, including those arising from parameterization dependence, have also been carefully considered. The resulting HERAPDFs (called HERAPDF0.1) have improved precision at low- $x$  compared to the global fits. this will be important for predictions of the  $W$  and  $Z$  cross-sections at the LHC, as explored in the next Section.

These PDFs have been released on LHAPDF in version LHAPDF.5.6: they consist of a central value and 22 experimental eigenvectors plus 12 model alternatives. The user should sum over  $N_{\text{mem}}=1,22$  for experimental uncertainties and over  $N_{\text{mem}}=1,34$  for total uncertainties.

### 1.6 Predictions for $W$ and $Z$ cross-sections at the LHC using the HERAPDF0.1

At leading order (LO),  $W$  and  $Z$  production occur by the process,  $q\bar{q} \rightarrow W/Z$ , and the momentum fractions of the partons participating in this subprocess are given by,  $x_{1,2} = \frac{M}{\sqrt{s}} \exp(\pm y)$ , where  $M$  is the centre of mass energy of the subprocess,  $M = M_W$  or  $M_Z$ ,  $\sqrt{s}$  is the centre of mass energy of the reaction ( $\sqrt{s} = 14$  TeV at the LHC) and  $y = \frac{1}{2} \ln \frac{(E+pl)}{(E-pl)}$  gives the parton

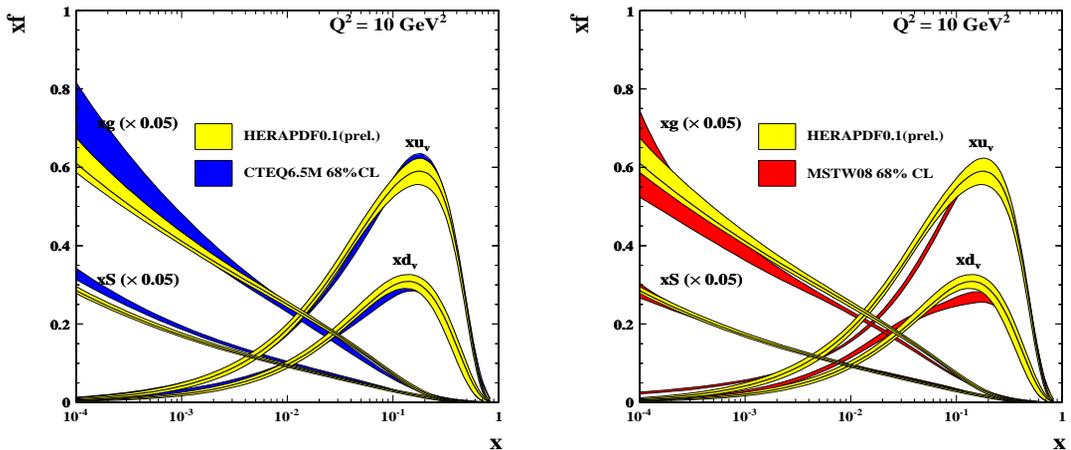


Fig. 14: HERAPDFs at  $Q^2 = 10\text{GeV}^2$  compared to the PDFs from CTEQ6.5 and MSTW08(prel.)

rapidity. The kinematic plane for LHC parton kinematics is shown in Fig. 15. Thus, at central rapidity, the participating partons have small momentum fractions,  $x \sim 0.005$ . Moving away from central rapidity sends one parton to lower  $x$  and one to higher  $x$ , but over the central rapidity range,  $|y| < 2.5$ ,  $x$  values remain in the range,  $5 \times 10^{-4} < x < 5 \times 10^{-2}$ . Thus, in contrast to the situation at the Tevatron, the scattering is happening mainly between sea quarks. Furthermore, the high scale of the process  $Q^2 = M^2 \sim 10,000 \text{ GeV}^2$  ensures that the gluon is the dominant parton, see Fig. 15, so that these sea quarks have mostly been generated by the flavour blind  $g \rightarrow q\bar{q}$  splitting process. Thus the precision of our knowledge of  $W$  and  $Z$  cross-sections at the LHC is crucially dependent on the uncertainty on the momentum distribution of the low- $x$  gluon.

HERA data have already dramatically improved our knowledge of the low- $x$  gluon, as discussed in earlier proceedings of the HERALHC workshop [13]. Now that the precision of HERA data at small- $x$  have been dramatically improved by the combination of H1 and ZEUS HERA-I data, we re-investigate the consequences for predictions of  $W, Z$  production at the LHC.

Predictions for the  $W/Z$  cross-sections, decaying to the lepton decay mode, using CTEQ, ZEUS PDFs and the HERAPDF0.1 are summarised in Table 2. Note that the uncertainties of CTEQ PDFs have been rescaled to represent 68% CL, in order to be comparable to the HERA PDF uncertainties. The precision on the predictions of the global fits (CTEQ6.1/5 and ZEUS-2002) for the total  $W/Z$  cross-sections is  $\sim 3\%$  at 68% CL. The precision of the ZEUS-2005 PDF fit prediction, which used only ZEUS data, is comparable, since information on the low- $x$  gluon is coming from HERA data alone. The increased precision of the HERAPDF0.1 low- $x$  gluon PDF results in increased precision of the  $W/Z$  cross-section predictions of  $\sim 1\%$ .

It is interesting to consider the predictions as a function of rapidity. Fig 16 shows the predictions for  $W^+, W^-, Z$  production as a function of rapidity from the HERAPDF0.1 PDF fit and compares them to the predictions from a PDF fit, using the same parameterization and assumptions, to the H1 and ZEUS data from HERA-I uncombined. The increase precision due

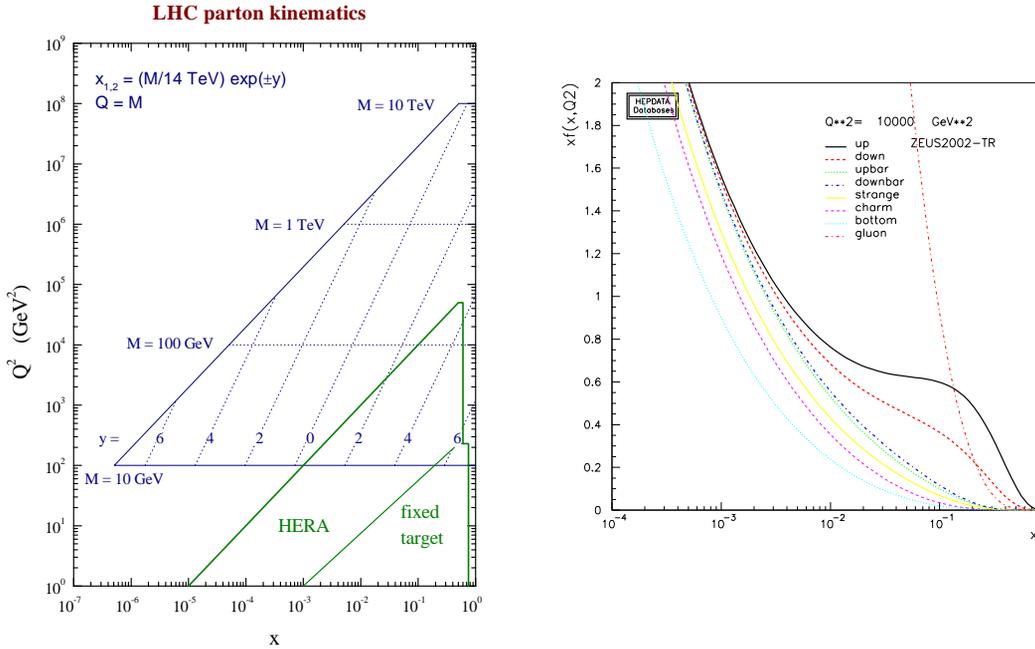


Fig. 15: Left plot: The LHC kinematic plane (thanks to James Stirling). Right plot: Typical PDF distributions at  $Q^2 = 10,000 \text{ GeV}^2$ .

to the combination is impressive. Fig. 17 show the predictions for  $W^+$ ,  $W^-$ ,  $Z$  production as a function of rapidity from the CTEQ6.1, 6.6 and MRST01 PDF fits for comparison. The uncertainties on the CTEQ and MRST PDF predictions have been rescaled to represent 68% CL limits, for direct comparability to the HERAPDF0.1 uncertainties. At central rapidity these limits give an uncertainty on the boson cross-sections of  $\sim 5\%$ , ( $\sim 3\%$ ), ( $\sim 2\%$ ) for CTEQ6.1, (CTEQ6.6), (MRST01) compared to  $\sim 1\%$  for the HERAPDF0.1.

So far, only experimental uncertainties have been included in these evaluations. It is also necessary to include model uncertainties. Fig. 18 shows the  $W^+$ ,  $W^-$ ,  $Z$  rapidity distributions including the six sources of model uncertainty detailed in Section 1.3. These model uncertainties increase the total uncertainty at central rapidity to  $\sim 2\%$ . Further uncertainty due to the choice of  $\alpha_s(M_Z)$  is small because, although a lower (higher) choice results in a larger (smaller) gluon at low  $x$ , the rate of QCD evolution is lower (higher) and this largely compensates. Uncertainties due to the choice of parameterization also have little impact on the boson rapidity spectra in the central region as illustrated in Fig. 18 by the superimposed blue line, which represents the alternative 'humpy' gluon parameterization (see Sec. 1.4).

Since the PDF uncertainty feeding into the  $W^+$ ,  $W^-$  and  $Z$  production is mostly coming from the gluon PDF, for all three processes, there is a strong correlation in their uncertainties, which can be removed by taking ratios. Figs. 16, 17 and 18 also show the  $W$  asymmetry

$$A_W = (W^+ - W^-)/(W^+ + W^-).$$

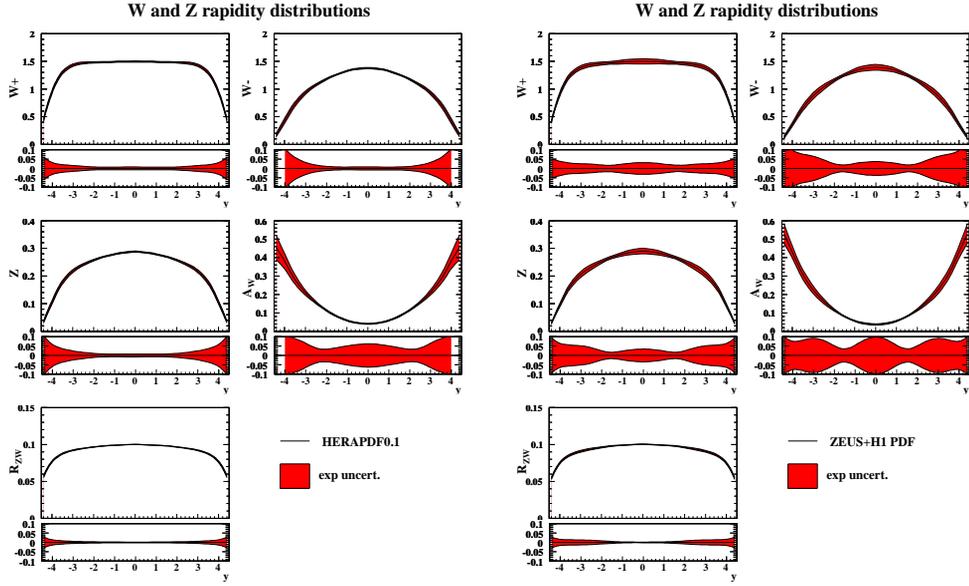


Fig. 16: The  $W^+$ ,  $W^-$ ,  $Z$  rapidity distributions,  $A_W$  and  $R_{ZW}$  (see text) and their uncertainties as predicted by (left) HERAPDF0.1 (right) a similar fit to the uncombined ZEUS and H1 data from HERA-I.

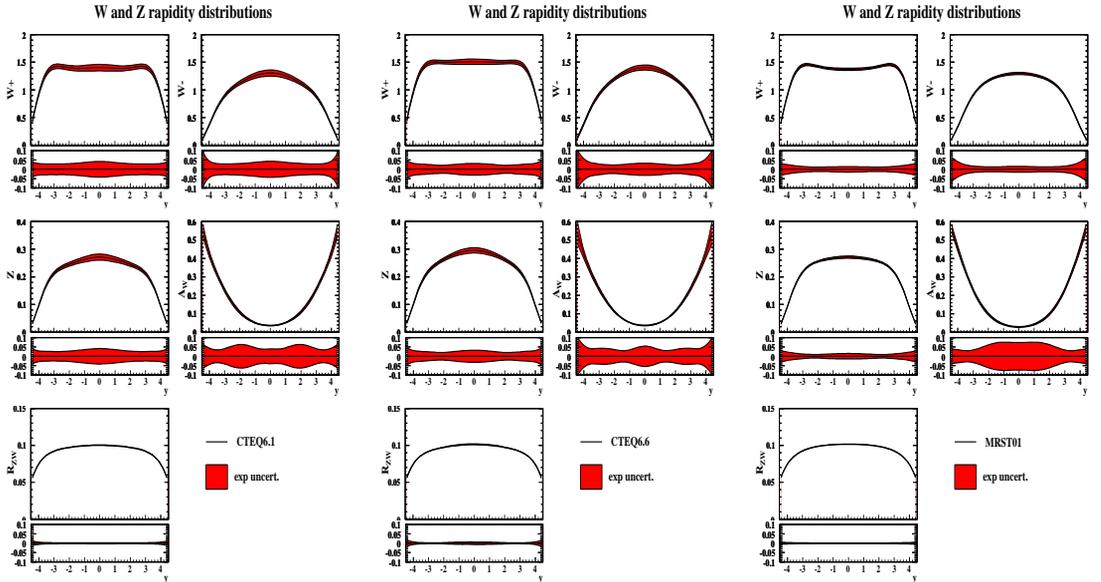


Fig. 17: The  $W^+$ ,  $W^-$ ,  $Z$  rapidity distributions,  $A_W$  and  $R_{ZW}$  (see text) and their uncertainties (scaled to 68% CL) as predicted by (left) CTEQ6.1, (middle) CTEQ6.6, right (MRST01)

PDF Set	$\sigma(W^+).B(W^+ \rightarrow l^+\nu_l)$	$\sigma(W^-).B(W^- \rightarrow l^-\bar{\nu}_l)$	$\sigma(Z).B(Z \rightarrow l^+l^-)$
CTEQ6.1	$11.61 \pm 0.34$ nb	$8.54 \pm 0.26$ nb	$1.89 \pm 0.05$ nb
CTEQ6.5	$12.47 \pm 0.28$ nb	$9.14 \pm 0.22$ nb	$2.03 \pm 0.04$ nb
ZEUS-2002	$12.07 \pm 0.41$ nb	$8.76 \pm 0.30$ nb	$1.89 \pm 0.06$ nb
ZEUS-2005	$11.87 \pm 0.45$ nb	$8.74 \pm 0.31$ nb	$1.97 \pm 0.06$ nb
HERAPDF0.1	$12.14 \pm 0.13$ nb	$9.08 \pm 0.14$ nb	$1.99 \pm 0.025$ nb

 Table 2: LHC  $W/Z$  cross-sections for decay via the lepton mode, for various PDFs, with 68% CL uncertainties.

The experimental PDF uncertainty on the asymmetry is larger ( $\sim 5\%$  for both CTEQ and HERAPDFs,  $\sim 7\%$  for the MRST01 PDFs) than that on the individual distributions and the variation between PDF sets is also larger - compare the central values of the CTEQ and MRST predictions, which are almost 25% discrepant. This is because the asymmetry is sensitive to the difference in the valence PDFs,  $u_v - d_v$ , in the low- $x$  region,  $5 \times 10^{-4} < x < 5 \times 10^{-2}$ , where there is no constraint from current data. To see this consider that at LO,

$$A_W \sim (u\bar{d} - d\bar{u})/(u\bar{d} + d\bar{u} + c\bar{s} + s\bar{c})$$

and that  $\bar{d} \sim \bar{u}$  at low- $x$ . (Note that the  $c\bar{s}$  and  $s\bar{c}$  contributions cancel out in the numerator). The discrepancy between the CTEQ and MRST01 asymmetry predictions at  $y = 0$  can be quantitatively understood by considering their different valence PDFs (see Figs. 13, 14 in Sec. 1.4). In fact a measurement of the asymmetry at the LHC will provide new information to constrain these PDFs.

By contrast, the ratio

$$R_{ZW} = Z/(W^+ + W^-),$$

also shown in Figs. 16, 17 and 18, has very small PDF uncertainties (both experimental and model) and there is no significant variation between PDF sets. To understand this consider that at LO

$$R_{ZW} = (u\bar{u} + d\bar{d} + c\bar{c} + s\bar{s})/(u\bar{d} + d\bar{u} + c\bar{s} + s\bar{c})$$

(modulo electroweak couplings) and that  $\bar{d} \sim \bar{u}$  at low- $x$ <sup>3</sup>. This will be a crucial measurement for our understanding of Standard Model Physics at the LHC.

However, whereas the  $Z$  rapidity distribution can be fully reconstructed from its decay leptons, this is not possible for the  $W$  rapidity distribution, because the leptonic decay channels which we use to identify the  $W$ 's have missing neutrinos. Thus we actually measure the  $W$ 's decay lepton rapidity spectra rather than the  $W$  rapidity spectra. Fig. 18 also shows the rapidity spectra for positive and negative leptons from  $W^+$  and  $W^-$  decay, the lepton asymmetry,

$$A_l = (l^+ - l^-)/(l^+ + l^-)$$

and the ratio

$$R_{Zl} = Z/(l^+ + l^-)$$

<sup>3</sup>There is some small model dependence from the strange sea fraction accounted for in both HERAPDF0.1 and in CTEQ6.6 PDFs.

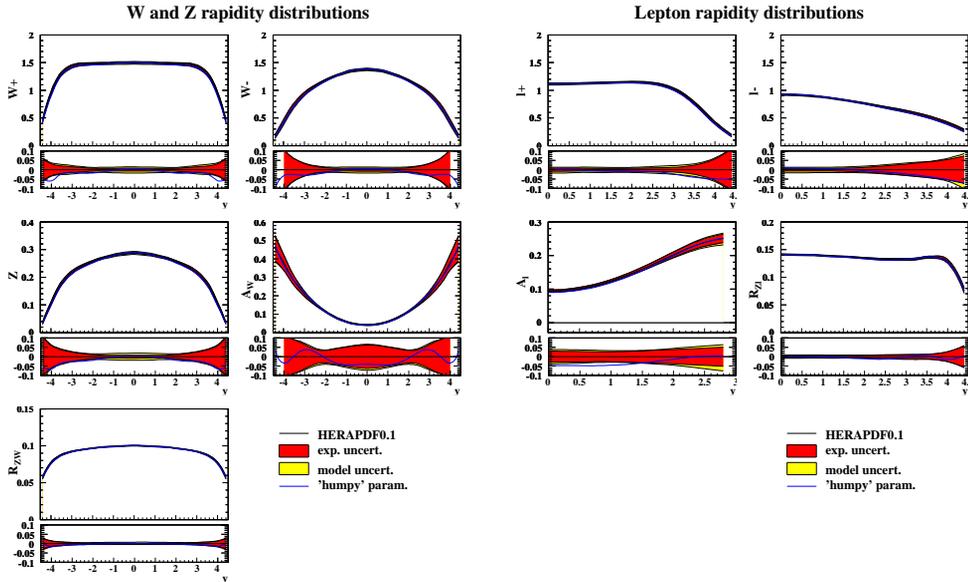


Fig. 18: Left: the  $W^+$ ,  $W^-$ ,  $Z$  rapidity distributions,  $A_W$ , and  $R_{ZW}$  (see text) and their experimental uncertainties (red) and model uncertainties (yellow). Right: the  $l^+$ ,  $l^-$  rapidity distributions,  $A_l$  and  $R_{Zl}$  (see text) and their experimental and model uncertainties. The superimposed blue line represents the results of the alternative 'humpy' gluon parameterization.

A cut of,  $p_{tl} > 25$  GeV, has been applied on the decay lepton, since it will not be possible to trigger on leptons with small  $p_{tl}$ . A particular lepton rapidity can be fed from a range of  $W$  rapidities so that the contributions of partons at different  $x$  values is smeared out in the lepton spectra, but the broad features of the  $W$  spectra remain.

In summary, these investigations indicate that PDF uncertainties, deriving from experimental error, on predictions for the  $W$ ,  $Z$  rapidity spectra in the central region, have reached a precision of  $\sim 1\%$ , due to the input of the combined HERA-I data. This level of precision is maintained when using the leptons from the  $W$  decay and gives us hope that we could use these processes as luminosity monitors<sup>4</sup>. However, model dependent uncertainties must now be considered very carefully. The current study will be repeated using a general-mass variable-flavour scheme for heavy quarks.

The predicted precision on the ratios  $R_{ZW}$ ,  $R_{Zl}$  is even better since model uncertainties are also very small giving a total uncertainty of  $\sim 1\%$ . This measurement may be used as a SM benchmark. However the  $W$  and lepton asymmetries have larger uncertainties (5 – 7%). A measurement of these quantities would give new information on valence distributions at small- $x$ .

<sup>4</sup>A caveat is that the current study has been performed using PDF sets which are extracted using NLO QCD in the DGLAP formalism. The extension to NNLO gives small corrections  $\sim 1\%$ . However, there may be much larger uncertainties in the theoretical calculations because the kinematic region involves low- $x$ . There may be a need to account for  $\ln(1/x)$  resummation or high gluon density effects.

## 2 Measurements of the Proton Structure Function $F_L$ at HERA <sup>5</sup>

### 2.1 Introduction

The inclusive deep inelastic  $ep$  scattering (DIS) cross section can at low  $Q^2$  be written in terms of the two structure functions,  $F_2$  and  $F_L$ , in reduced form as

$$\sigma_r(x, Q^2, y) \equiv \frac{d^2\sigma}{dx dQ^2} \cdot \frac{Q^4 x}{2\pi\alpha^2 Y_+} = F_2(x, Q^2) - \frac{y^2}{Y_+} \cdot F_L(x, Q^2), \quad (4)$$

where  $Q^2 = -q^2$  is the negative of the square of the four-momentum transferred between the electron<sup>6</sup> and the proton, and  $x = Q^2/2qP$  denotes the Bjorken variable, where  $P$  is the four-momentum of the proton. The two variables are related through the inelasticity of the scattering process,  $y = Q^2/sx$ , where  $s = 4E_e E_p$  is the centre-of-mass energy squared determined from the electron and proton beam energies,  $E_e$  and  $E_p$ . In eq. 4,  $\alpha$  denotes the fine structure constant and  $Y_+ = 1 + (1 - y)^2$ .

The two proton structure functions  $F_2$  and  $F_L$  are related to the cross sections of the transversely and longitudinally polarised virtual photons interacting with protons,  $\sigma_L$  and  $\sigma_T$ , according to  $F_L \propto \sigma_L$  and  $F_2 \propto (\sigma_L + \sigma_T)$ . Therefore the relation  $0 \leq F_L \leq F_2$  holds. In the Quark Parton Model (QPM),  $F_2$  is the sum of the quark and anti-quark  $x$  distributions, weighted by the square of the electric quark charges, whereas the value of  $F_L$  is zero [14]. The latter follows from the fact that a quark with spin  $\frac{1}{2}$  cannot absorb a longitudinally polarised photon.

In Quantum Chromodynamics (QCD),  $F_L$  differs from zero, receiving contributions from quarks and from gluons [15]. At low  $x$  and in the  $Q^2$  region of deep inelastic scattering the gluon contribution greatly exceeds the quark contribution. Therefore  $F_L$  is a direct measure of the gluon distribution to a very good approximation. The gluon distribution is also constrained by the scaling violations of  $F_2$  as described by the DGLAP QCD evolution equations [7–9, 16, 17]. An independent measurement of  $F_L$  at HERA, and its comparison with predictions derived from the gluon distribution extracted from the  $Q^2$  evolution of  $F_2(x, Q^2)$ , thus represents a crucial test on the validity of perturbative QCD (pQCD) at low  $x$ . Moreover, depending on the particular theoretical approach adopted, whether it be a fixed order pQCD calculation, a resummation scheme, or a color dipole ansatz, there appear to be significant differences in the predicted magnitude of  $F_L$  at low  $Q^2$ . A measurement of  $F_L$  may be able to distinguish between these approaches.

Previously the structure function  $F_L$  was extracted by the H1 collaboration from inclusive data at high  $y$  using indirect methods, as discussed in Sect. 2.2. A preliminary measurement was also presented by the ZEUS collaboration using initial state radiation (ISR) events [18], although the precision of this measurement was limited.

To make a direct measurement of  $F_L$ , reduced cross sections must be measured at the same  $x$  and  $Q^2$  but with different  $y$  values. This can be seen from eq. 4 which states that  $F_L(x, Q^2)$  is equal to the partial derivative  $\partial\sigma_r(x, Q^2, y)/\partial(y^2/Y_+)$ . Due to the relationship  $y = Q^2/xs$  this requires data to be collected at different beam-beam centre-of-mass energies, which was done in the last year of HERA running. To maximize the precision of this procedure, the measurable

<sup>5</sup>Contributing authors: J. Grebenyuk, V. Lendermann

<sup>6</sup>The term electron is used here to denote both electrons and positrons unless the charge state is specified explicitly.

range of  $y^2/Y_+$  had to be maximised for each fixed  $x$  and  $Q^2$ . This was achieved by operating HERA at the lowest attainable centre-of-mass energy and by measuring this data up to the highest possible value of  $y$ . An intermediate HERA centre-of-mass energy was also chosen, to improve the precision of  $F_L$  extraction and to act as a consistency check. More specifically, between March and June 2007, HERA was operated with proton beam energies,  $E_p = 460$  GeV and 575 GeV, compared to the previous nominal value of 920 GeV. The electron beam energy was unaltered at  $E_e = 27.6$  GeV. Thus, three data sets, referred to the high- (HER), middle- (MER) and low-energy running (LER) samples, were collected with  $\sqrt{s} = 318$  GeV, 251 GeV and 225 GeV, respectively. The integrated luminosities of the data sets used by ZEUS (H1) to measure  $F_L$  are  $32.8$  ( $21.6$ )  $\text{pb}^{-1}$  for HER,  $6$  ( $6.2$ )  $\text{pb}^{-1}$  for MER and  $14$  ( $12.4$ )  $\text{pb}^{-1}$  for LER. The specific issues of the recent H1 and ZEUS analyses are discussed in Sect. 2.3, and the results are presented in Sect. 2.4.

## 2.2 Indirect $F_L$ Extraction by H1

H1 extracted  $F_L$  from inclusive data using several indirect methods, which exploit the turn over of the reduced cross section at high  $y$  due to the  $F_L$  contribution. The basic principle is the following. First, the reduced neutral current cross section  $\sigma_r$  is measured in a  $y$  range, where the  $F_L$  contribution is negligible and thus the relation  $\sigma_r = F_2$  holds very well. Afterward, based on some theoretical assumption, the knowledge of  $F_2$  is extrapolated towards high  $y$ . Finally  $F_L$  is extracted from the difference between the prediction for  $F_2$  and the measurement of  $\sigma_r$  at high  $y$ .

In the analyses at  $Q^2 \gtrsim 10 \text{ GeV}^2$  [4, 19, 20] the ‘‘extrapolation’’ method is used. In this method, an NLO QCD PDF fit to H1 HERA I data is performed at  $y < 0.35$ , and the results are extrapolated to higher  $y$  using the DGLAP evolution equations.  $F_L$  is then extracted at a fixed  $y = 0.75$  and at  $Q^2$  up to  $700 \text{ GeV}^2$  using eq. 4. The extracted values are shown in Fig. 19 for the high- $Q^2$  analysis [4].

At low  $Q^2$ , extrapolations of DGLAP fits become uncertain. For  $Q^2 \lesssim 2 \text{ GeV}^2$ , as the strong coupling constant  $\alpha_s(Q^2)$  increases, the higher order corrections to the perturbative expansion become large and lead to the breakdown of the pQCD calculations. Therefore other methods are used in the H1 low- $Q^2$  data analyses.

The ‘‘shape method’’, as used in the last H1 low- $Q^2$  study of HERA I data [21], exploits the shape of  $\sigma_r$  in a given  $Q^2$  bin. The  $Q^2$  dependence at high  $y$  is driven by the kinematic factor  $y^2/Y_+$  (eq. 4), and to a lesser extent by  $F_L(x, Q^2)$ . On the other hand, the gluon dominance at low  $x$  suggests that  $F_L$  may exhibit an  $x$  dependence similar to  $F_2$ . Therefore it is assumed that  $F_L$  is proportional to  $F_2$  and the coefficient of proportionality depends only on  $Q^2$ . In the extraction procedure one uses the ratio  $R$  of the cross sections of the transversely and longitudinally polarised photons

$$R = \frac{\sigma_T}{\sigma_L} = \frac{F_L}{F_2 - F_L} \quad (5)$$

which is thus assumed to depend only on  $Q^2$ . The reduced cross section is fitted by

$$\sigma_r = F_2 \left[ 1 - \frac{y^2}{Y_+} \frac{R(Q^2)}{1 + R(Q^2)} \right], \quad (6)$$

where some phenomenological model for  $F_2$  is chosen.

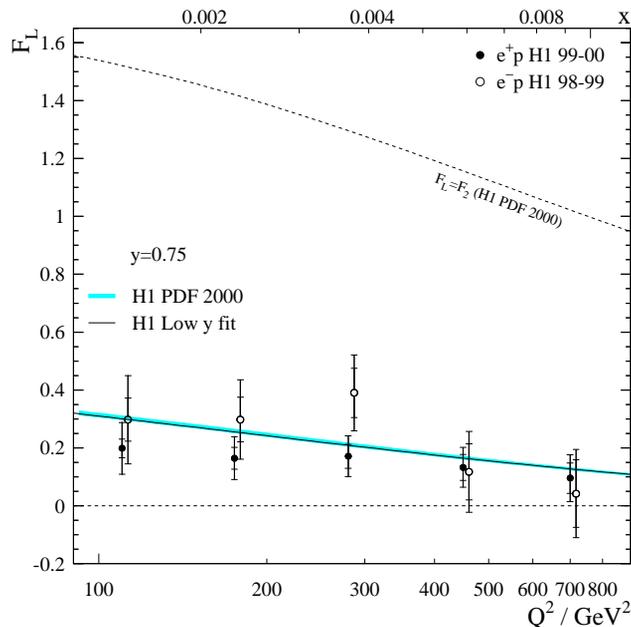


Fig. 19:  $F_L$  determined indirectly by H1 at a fixed  $y = 0.75$  and high  $Q^2$  is shown as a function of  $Q^2$  (lower scale) or equivalently  $x$  (upper scale) for  $e^+p$  (closed circles) and  $e^-p$  (open circles) data. The inner error bar represents the statistical error, and the outer error bar also includes the systematic error and the uncertainty arising from the extrapolation of  $F_2$ .

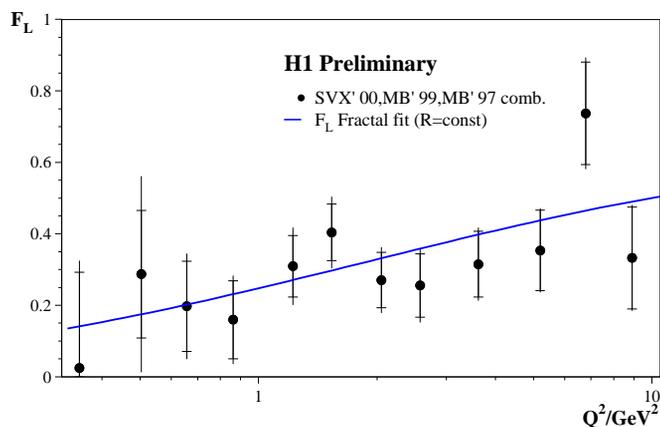


Fig. 20:  $Q^2$  dependence of  $F_L(x, Q^2)$  at fixed  $y = 0.75$ , extracted from the preliminary H1 low- $Q^2$  data. The solid line shows the prediction of the fractal fit with a constant  $R$ .

An example of such an extraction using a fractal fit for  $F_2$  [22] is shown in Fig. 20, where preliminary H1 results [21] for  $F_L$  at  $y = 0.75$  in the range of  $0.35 \leq Q^2 \leq 8.5 GeV^2$  are presented. The data favour a positive, not small  $F_L$  at low  $Q^2$ . A drawback of this method is that it reveals a considerable dependence of  $R$  on the choice of the  $F_2$  model.

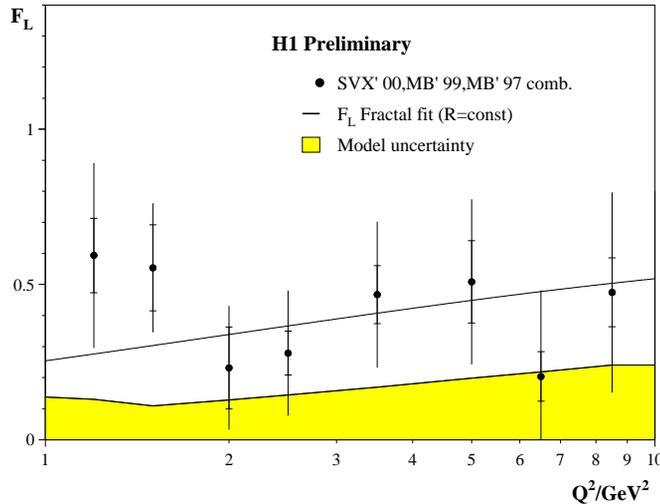


Fig. 21: Structure function  $F_L$  extracted by H1 using the derivative method. The solid line shows the prediction of the fractal fit with a constant  $R$ . The inner error bars represent statistical uncertainties, the outer error bars represent statistical and systematic uncertainties added in quadrature. The solid (yellow) band indicates the model uncertainty.

In the derivative method [20, 21],  $F_L$  is extracted from the partial derivative of the reduced cross section on  $y$  at fixed  $Q^2$

$$\left. \frac{\partial \sigma_r}{\partial \ln y} \right|_{Q^2} = -x \frac{\partial F_2}{\partial x} - \frac{2y^2(2-y)}{Y_+^2} F_L - x \frac{y^2}{Y_+} \frac{\partial F_L}{\partial x} \quad (7)$$

which is dominated by the  $F_L$ -dependent term at high  $y$ . The term proportional to  $\partial F_L / \partial x$  is negligible for moderately varying parametrisations of  $F_L$ . For low  $Q^2$  values the rise of  $F_2$  is weak. The change of the term  $x \partial F_2 / \partial x$  for the two assumptions: no rise at low  $x$ , i.e.  $\partial F_2 / \partial x = 0$ , and  $F_2 \propto x^{-\lambda}$  is numerically significantly smaller than the experimental precision for  $\partial \sigma_r / \partial \ln y$ . Therefore the derivative method provides a means for determining  $F_L$  at low  $Q^2$  with minimal phenomenological assumption. On the other hand, the errors obtained with the derivative method turn out to be significantly larger than those from the shape method.

The preliminary results of  $F_L$  extraction from H1 HERAI data [21] are presented in Fig. 21. The residual dependence of the measurement on the assumption made for  $F_2$  is estimated by a comparison with results obtained assuming an  $F_2$  which is flat in  $y$ . The lower bound on  $F_L$  obtained this way is depicted as a solid band in the figure.

### 2.3 Details of Direct $F_L$ Measurements

The H1 and ZEUS analysis procedures involve a measurement of the inclusive cross section at  $y > 0.1$ . In this range, the kinematic variables  $x$ ,  $y$  and  $Q^2$  are most accurately reconstructed using the polar angle,  $\theta_e$ , and the energy,  $E'_e$ , of the scattered electron according to

$$y = 1 - \frac{E'_e \sin^2 \theta_e}{E_e}, \quad Q^2 = \frac{E_e'^2 \sin^2 \theta_e}{1 - y}, \quad x = \frac{Q^2}{y s}. \quad (8)$$

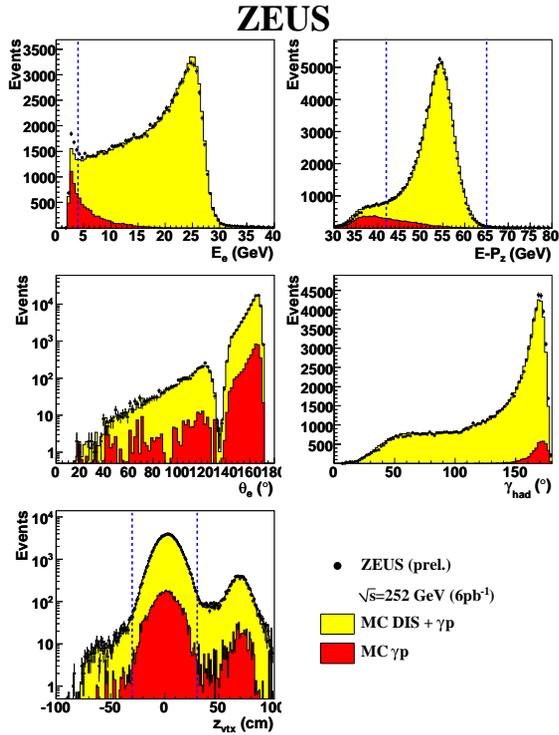


Fig. 22: Comparison of 575 GeV data with the sum of DIS and background simulations for the energy of the scattered electron, total  $E - p_z$ , theta of the scattered electron, angle of the hadronic final state and  $z$  coordinate of the vertex. The dotted lines indicate the cuts applied.

Reaching the high  $y$  values necessary for the  $F_L$  determination requires a measurement of the scattered electron with energy down to a few GeV. The electron candidate is selected as an isolated electromagnetic energy deposition (cluster) in a calorimeter. The crucial analysis issue at high- $y$  region is the identification of the scattered electron, and the estimation of the hadronic background which occurs when a particle from the hadronic final state mimics the electron signal. Most of background events are photoproduction ( $\gamma p$ ) events with  $Q^2 \approx 0$  in which the final state electron is scattered at low angles (high  $\theta$ )<sup>7</sup> and thus escapes through the beam pipe.

The  $\gamma p$  background suppression is performed in several steps. Firstly, calorimeter shower estimators are utilised which exploit the different profiles of electromagnetic and hadronic showers. Secondly, background coming from neutral particles, such as  $\pi_0$ , can be rejected by requiring a track associated to the electron candidate. Furthermore,  $\gamma p$  events are suppressed by utilising the energy-momentum conservation. For that, the variable  $E - p_z = \sum_i (E_i - p_{z,i})$  is exploited, where the sum runs over energies  $E_i$  and longitudinal momentum components  $p_{z,i}$  of all particles in the final state. The requirement  $E - p_z > 35$  (42) GeV in the H1 (ZEUS) analysis removes

<sup>7</sup>The  $z$  axis of the right-handed coordinate systems used by H1 and ZEUS is defined by the direction of the incident proton beam with the origin at the nominal  $ep$  interaction vertex. Consequently, small scattering angles of the final state particles correspond to large polar angles in the coordinate system.

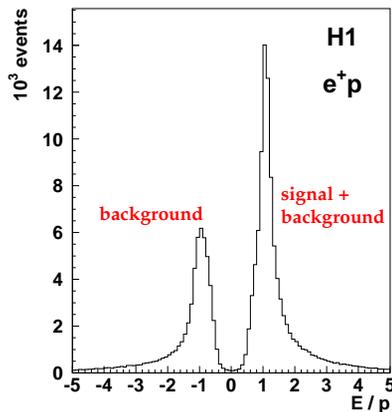


Fig. 23: Distribution of energy over momentum for tracks linked to clusters in the SpaCal with energy from 3.4 to 10 GeV that pass all the medium  $Q^2$  analysis cuts. Tracks with a negative charge are assigned a negative  $E/p$ .

events where the escaping electron carries a significant momentum. It also suppresses events with hard initial state photon radiation.

However, at low  $E'_e$  the remaining background contribution after such a selection is of a size comparable to or even exceeding the genuine DIS signal. The further analysis steps differ for the H1 and ZEUS analyses as discussed in the following.

**ZEUS Analysis Procedure** The electron candidates are selected as compact electromagnetic energy depositions in the Uranium Calorimeter (UCal). The position of the candidate is reconstructed using either the Small Angle Rear Tracking Detector (SRTD), which is a high-granularity lead-scintillator calorimeter, or with the Hadron-Electron Separator (HES), which is a silicon detector located in the electromagnetic section of the UCal. The candidates are selected such that  $E'_e > 6 \text{ GeV}$ <sup>8</sup>.

The candidates are validated using information from the tracking devices. The acceptance region for ZEUS tracking is limited to polar angles  $\theta_e \lesssim 154^\circ$ . The tracking detectors do provide some coverage beyond  $\theta_e = 154^\circ$ , up to  $\theta_e \approx 168^\circ$ , however the number of tracking layers is too sparse for full track reconstruction. The hit information from the tracking detectors can still be used. To do this, a “road” is created between the measured interaction vertex and the position of the electron candidate in the calorimeter. Hits in the tracking layers along the road are then counted and compared to the maximum possible number of hits. If too few hits are found, the candidate is assumed to be a neutral particle and it is rejected. To ensure the reliability of this method, the scattered electron is required to exit the central drift chamber at a radius  $R > 20 \text{ cm}$ . Given that  $E'_e > 6 \text{ GeV}$ , this effectively limits the maximal  $y$  to  $y < 0.8$  and the minimum  $Q^2$  achievable at low  $y$ . In the HES analysis, events are measured down to  $y = 0.2$  roughly translating to the  $Q^2$  region,  $Q^2 > 24 \text{ GeV}^2$ . No background treatment based on the charge of the candidate is performed.

<sup>8</sup>Cut of  $E'_e > 4 \text{ GeV}$  is used for the event selection, although the binning for  $F_L$  measurement is chosen such that  $E'_e > 6 \text{ GeV}$ .

The remaining  $\gamma p$  background is estimated using Monte Carlo (MC) simulations. In order to minimise the model uncertainty of the  $\gamma p$  simulation, a pure photoproduction sample is selected using an electron tagger placed close to the beam pipe about 6 meters away from the interaction point in the rear direction. It tags, with almost perfect efficiency and purity, the scattered electrons in such events which are not identified in the main detector and escape down the beam pipe. Photoproduction MC is verified against and normalised to this sample. The normalisation factor is found to be  $1 \pm 0.1$  for all data sets.

Figure 22 shows, as an example, comparisons of the 575 GeV data with simulated distributions, for the energy of the scattered electron, total  $E - p_z$ , polar angle of the scattered electron, angle of the hadronic final state and the  $z$  coordinate of the interaction vertex. A good description of the data by the simulation is observed. A similar level of agreement was found for both, HER and LER data sets.

A full set of systematic uncertainties is evaluated for the cross section measurements. The largest single contribution comes from the electron energy scale uncertainty, which is known to within  $\pm 1\%$  for  $E'_e > 10$  GeV, increasing to  $\pm 3\%$  at  $E'_e = 5$  GeV. Other significant contributions are due to the  $\pm 10\%$  uncertainty in verifying the Pythia prediction of the  $\gamma p$  cross section using the electron tagger. The systematic uncertainty due to the luminosity measurement was reduced by scaling the three cross sections relative to each other. The spread of relative normalisation factor was found to be within the expected level of uncorrelated systematic uncertainty.

**H1 Analysis Procedure** The H1 measurements of  $F_L$  are performed in separate analyses involving different detector components and thus covering different  $Q^2$  ranges. In the high- $Q^2$  analysis the electron candidate is selected as an isolated electromagnetic energy deposition in the Liquid Argon (LAr) calorimeter which covers the polar angle range  $4^\circ < \theta < 153^\circ$ . The selected cluster is further validated by a matching track reconstructed in the central tracking device (CT) with an angular acceptance of  $15^\circ < \theta < 165^\circ$ . In the medium  $Q^2$  analysis the electron candidate is selected in the backward calorimeter SpaCal covering the angular range  $153^\circ < \theta < 177.5^\circ$  and is also validated by a CT track. Lower  $Q^2$  values are expected to be accessed in the third analysis, in which the SpaCal cluster is validated by a track in the Backward Silicon Tracker reaching the highest  $\theta$ . The first measurement of  $F_L$  at medium  $Q^2$  is already published [23], and preliminary results of the combined medium-high- $Q^2$  analysis are available.

The remaining  $\gamma p$  background is subtracted on statistical basis. The method of background subtraction relies on the determination of the electric charge of the electron candidate from the curvature of the associated track.

Figure 23 shows the  $E/p$  distribution of the scattered electron candidates from  $e^+p$  interactions with the energy  $E$  measured in the SpaCal and the momentum  $p$  of the linked track determined by the CT. The good momentum resolution leads to a clear distinction between the negative and positive charge distributions. The smaller peak corresponds to tracks with negative charge and thus represents almost pure background. These tracks are termed wrong sign tracks and events with such candidates are rejected. The higher peak, due to right sign tracks, contains the genuine DIS signal superimposed on the remaining positive background. The size of the latter to first approximation equals the wrong sign background. The principal method of background

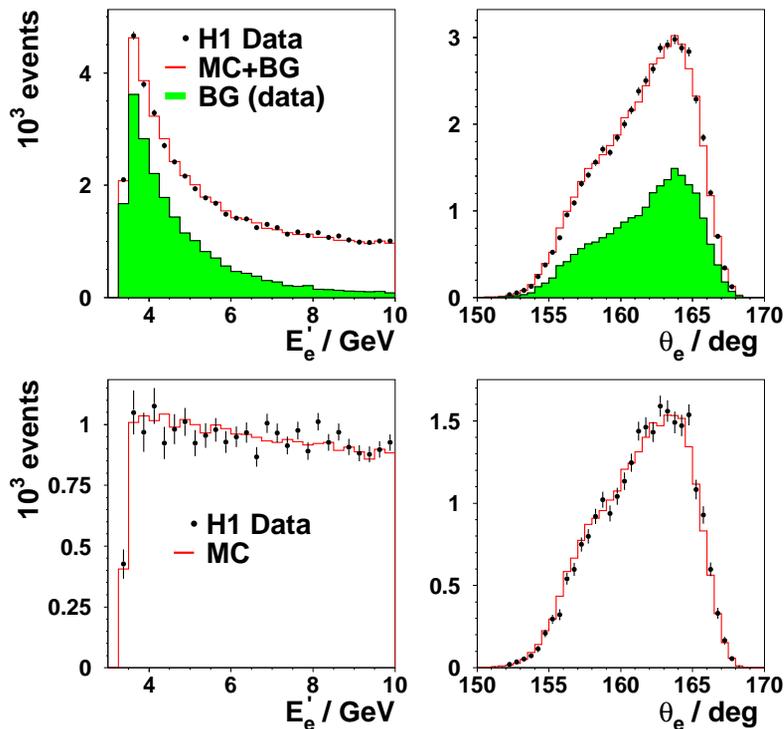


Fig. 24: Top: comparison of the correct sign data (points) with the sum (open histogram) of the DIS MC simulation and background, determined from the wrong sign data (shaded histogram), for the energy  $E'_e$  (left) and the polar angle  $\theta_e$  (right) of the scattered electron, for the 460 GeV data with  $E'_e < 10$  GeV. Bottom: as top but after background subtraction.

subtraction, and thus of measuring the DIS cross section up to  $y \simeq 0.9$ , consists of the subtraction of the wrong sign from the right sign event distribution in each  $x, Q^2$  interval.

The background subtraction based on the charge measurement requires a correction for a small but non-negligible charge asymmetry in the negative and positive background samples, as has been observed previously by H1 [20]. The main cause for this asymmetry lies in the enhanced energy deposited by anti-protons compared to protons at low energies. The most precise measurement of the background charge asymmetry has been obtained from comparisons of samples of negative tracks in  $e^+p$  scattering with samples of positive tracks in  $e^-p$  scattering. An asymmetry ratio of negative to positive tracks of 1.06 is measured using the high statistics  $e^\pm p$  data collected by H1 in 2003-2006. This result is verified using photoproduction events with a scattered electron tagged in a subdetector of the luminosity system.

Figure 24 shows, as an example, comparisons of the 460 GeV high  $y$  data with simulated distributions, for the energy and the polar angle of the scattered electron prior to and after subtraction of the background, which is determined using wrong sign data events.

The measurement of  $F_L$  as described below relies on an accurate determination of the variation of the cross section for a given  $x$  and  $Q^2$  at different beam energies. In order to reduce

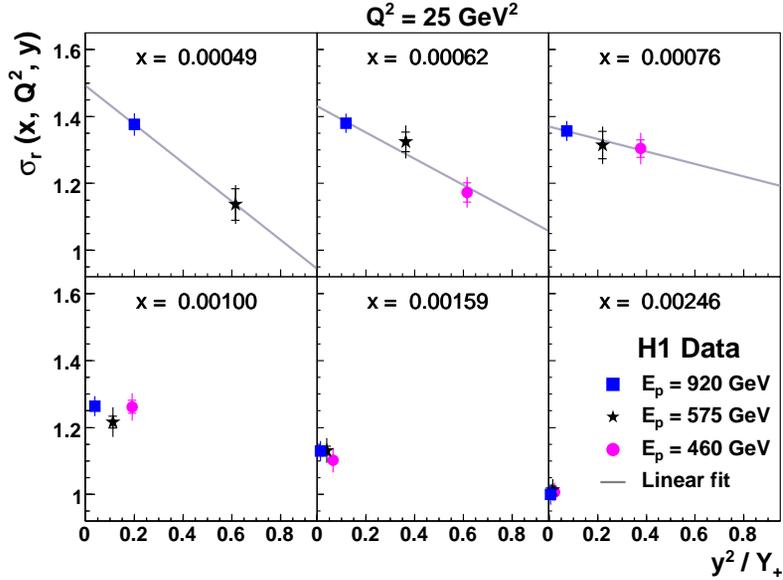


Fig. 25: The reduced inclusive DIS cross section plotted as a function of  $y^2/Y_+$  for six values of  $x$  at  $Q^2 = 25 \text{ GeV}^2$ , measured by H1 for proton beam energies of 920, 575 and 460 GeV. The inner error bars denote the statistical error, the full error bars include the systematic errors. The luminosity uncertainty is not included in the error bars. For the first three bins in  $x$ , corresponding to larger  $y$ , a straight line fit is shown, the slope of which determines  $F_L(x, Q^2)$ .

the uncertainty related to the luminosity measurement, which presently is known to 5% for each proton beam energy of the 2007 data, the three data samples are normalised relatively to each other. The renormalisation factors are determined at low  $y$ , where the cross section is determined by  $F_2$  only, apart from a small correction due to  $F_L$ . The relative normalisation is known to within 1.6%.

All correlated and uncorrelated systematic errors combined with the statistical error lead to an uncertainty on the measured cross sections at high  $y$  of 3 to 5%, excluding the common luminosity error.

#### 2.4 Measurements of $F_L(x, Q^2)$ by H1 and ZEUS

The longitudinal structure function is extracted from the measurements of the reduced cross section as the slope of  $\sigma_r$  versus  $y^2/Y_+$ , as can be seen in eq. 4. This procedure is illustrated in Fig. 25. The central  $F_L$  values are determined in straight-line fits to  $\sigma_r(x, Q^2, y)$  as a function of  $y^2/Y_+$  using the statistical and uncorrelated systematic errors.

The first published H1 measurement of  $F_L(x, Q^2)$  is shown in Fig. 26, the preliminary ZEUS measurement is presented in Fig. 27. The H1 measured values of  $F_L$  are compared with the H1 PDF 2000 fit [4], while the ZEUS  $F_L$  values are compared to the ZEUS-JETS PDF fit [3]. Both measurements are consistent and show a non-zero  $F_L$ .

The H1 results were further averaged over  $x$  at fixed  $Q^2$ , as shown in the left panel of Fig. 28. The averaging is performed taking the  $x$  dependent correlations between the systematic

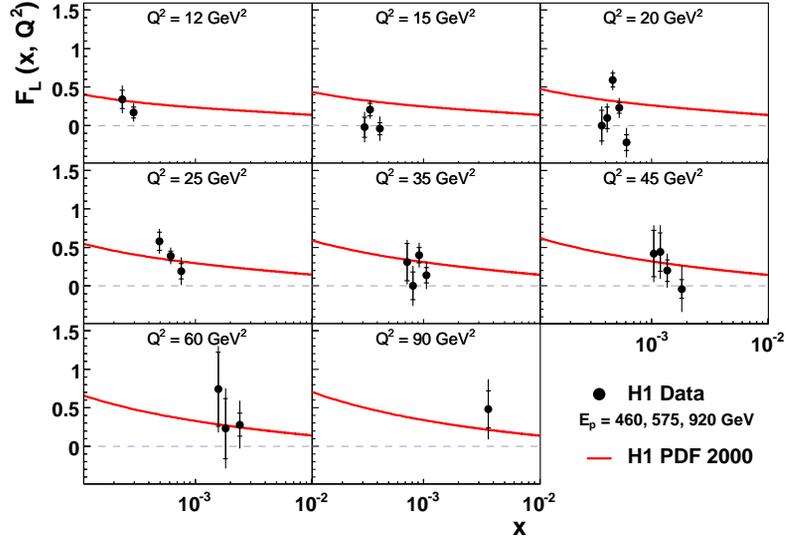


Fig. 26: The longitudinal proton structure function  $F_L(x, Q^2)$  measured by the H1 collaboration. The inner error bars denote the statistical error, the full error bars include the systematic errors. The curves represent the H1 PDF 2000 fit.

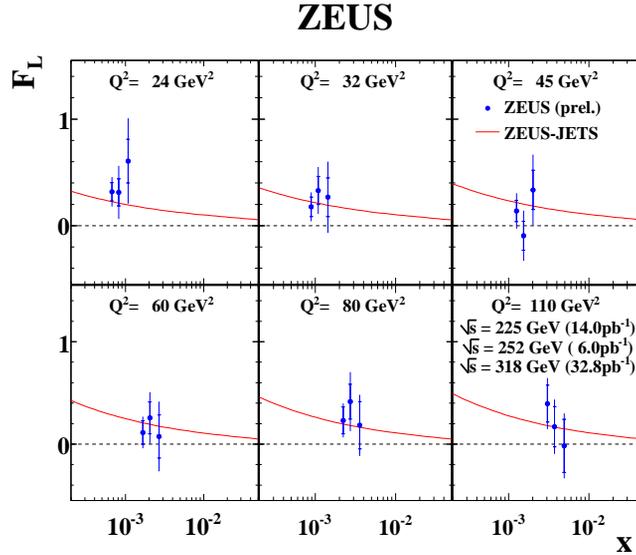


Fig. 27: The longitudinal proton structure function  $F_L(x, Q^2)$  measured by the ZEUS collaboration. The inner error bars denote the statistical error, the full error bars include the systematic errors. The curves represent the ZEUS-JETS PDF fit.

errors into account. The averaged values of  $F_L$  are compared with H1 PDF 2000 fit and with the expectations from global parton distribution fits at higher order perturbation theory performed by the MSTW [24] and the CTEQ [2, 25] groups. Within the experimental uncertainties the data are consistent with these predictions. The measurement is also consistent with previous indirect

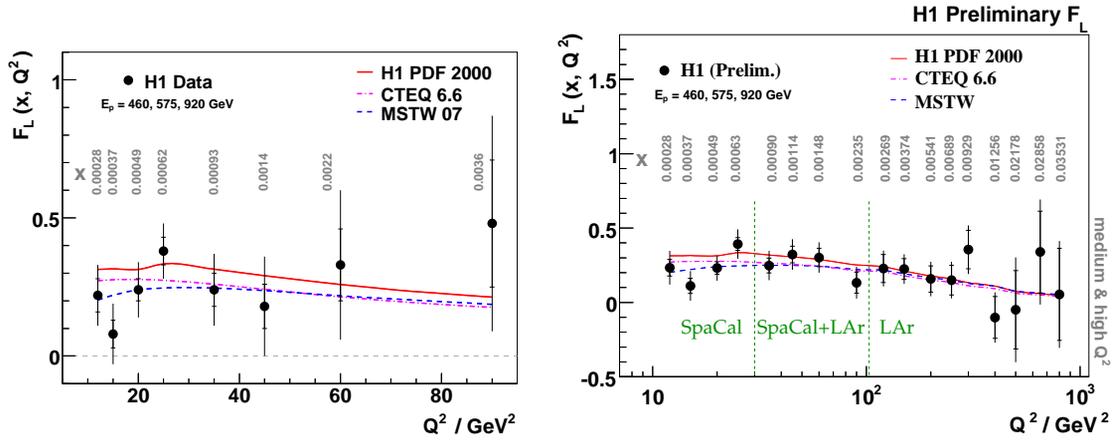


Fig. 28: The proton structure function  $F_L$  shown as a function of  $Q^2$  at the given values of  $x$ : a) first direct measurement at HERA by H1; b) preliminary H1 results combining SpaCal and LAr analyses. The inner error bars denote the statistical error, the full error bars include the systematic errors. The luminosity uncertainty is not included in the error bars. The solid curve describes the expectation on  $F_L$  from the H1 PDF 2000 fit using NLO QCD. The dashed (dashed-dotted) curve depicts the expectation of the MSTW (CTEQ) group using NNLO (NLO) QCD. The theory curves connect predictions at the given  $(x, Q^2)$  values by linear extrapolation.

determinations of  $F_L$  by H1.

In the combined medium–high  $Q^2$  analysis by H1 the  $Q^2$  range is extended up to  $Q^2 = 800 \text{ GeV}^2$ . The preliminary results are shown in the right panel of Fig. 28. In some  $Q^2$  bins there is an overlap between the SpaCal and LAr measurements which improves the precision of the  $F_L$  extraction as compared to the pure SpaCal analysis.

## 2.5 Summary

Direct measurements of the proton structure function  $F_L$  have been performed in deep inelastic  $ep$  scattering at low  $x$  at HERA. The  $F_L$  values are extracted by the H1 and ZEUS collaborations from the cross sections measured at fixed  $x$  and  $Q^2$  but different  $y$  values. This is achieved by using data sets collected with three different proton beam energies. The H1 and ZEUS results are consistent with each other and exhibit a non-zero  $F_L$ . The measurements are also consistent with the previous indirect determinations of  $F_L$  by H1. The results confirm DGLAP NLO and NNLO QCD predictions for  $F_L(x, Q^2)$ , derived from previous HERA data, which are dominated by a large gluon density at low  $x$ .

## References

- [1] A. M. et al, Eur. Phys.J C **23**, 73 (2002).
- [2] L. L. J. Pumplin, H. and K. T. W. Phys. Rev. D **75**, 054029 (2007). hep-ph/0701220.
- [3] ZEUS Collaboration, C. e. a. ZEUS Coll., S. Eur. Phys. J. C **42**, 1 (2005). hep-ph/0503274.

- [4] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C30**, 1 (2003). hep-ex/0304003.
- [5] A. Cooper-Sarkar, Phys. Rev **D 67**, 012007 (2003).
- [6] ZEUS and H. Collaborations, *Combination of h1 and zeus deep inelastic e+ - p scattering cross sections*. Preprint ZEUS-prel-07-036, H1prelim-07-007.
- [7] L. D. Yu. Sov. Phys. JETP **46**, 641 (1977).
- [8] N. G. V. and N. L. L. Sov. J. Nucl. Phys. **15**, 438 (1972).
- [9] G. Altarelli and G. Parisi, Nucl. Phys. **B 126**, 298 (1977).
- [10] M. Botje, *Qcdnum16.12*. Available from <http://www.nikhef.nl/h24/qcdnum/>.
- [11] C. Pascaud and F. Zomer, *Qcd analysis from the proton structure function f2 measurement: Issues on fitting, statistical and systematic errors*. Preprint LAL-95-05, 1995.
- [12] Particle Data Group Collaboration, C. Amsler *et al.*, Phys. Lett. **B667**, 1 (2008).
- [13] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.
- [14] C. Callan and D. Gross, Phys. Rev. Lett. **22**, 156 (1969).
- [15] F. W. A. Zee and B. T. S. Phys. Rev. **D 10**, 2881 (1974);  
G. Altarelli and G. Martinelli, Phys. Lett. **B 76**, 89 (1978).
- [16] N. G. V. and N. L. L. Sov. J. Nucl. Phys. **15**, 675 (1972).
- [17] N. L. L. Sov. J. Nucl. Phys. **20**, 94 (1975).
- [18] S. C. e. a. ZEUS Coll., in *EPS 2003 conference, Aachen*. 2003.
- [19] H1 Collaboration, C. A. e. a. H1 Coll., Phys. Lett. **B 393**, 452 (1997).  
hep-ex/9611017.
- [20] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C21**, 33 (2001). hep-ex/0012053.
- [21] A. V. Trevino, *Measurement of the Inclusive ep Scattering Cross Section at low Q<sup>2</sup> and x at HERA*, in *Proceedings of the 15th International Workshop on Deep-Inelastic Scattering*. 2007.
- [22] T. Lastovicka, Eur. Phys. J. **C 24**, 529 (2002). hep-ph/0203260.
- [23] H1 Collaboration, D. A. e. a. H1 Coll., Phys. Lett. **B 665**, 139 (2008). 0805.2809.
- [24] W. J. S. R. S. T. A. D. Martin and G. Watt, Phys. Lett. **B 652**, 292 (2007).  
hep-ph/0706.0459.
- [25] P. M. Nadolsky *et al.*, Phys. Rev. **D78**, 013004 (2008). 0802.0007.

# Proton - proton luminosity, standard candles and PDFs at the LHC

*J. Anderson, M. Boonekamp, H. Burkhardt, M. Dittmar, V. Halyo, T. Petersen*

## 1 Introduction

The Large Hadron Collider (LHC) is expected to start colliding proton beams in 2009, and is expected to reach design parameters in energy and luminosity sometime later and deliver a few  $fb^{-1}$  per year of data at the 14 TeV collision energy.

During the past 15 years many theoretical calculations and experimental simulations have demonstrated a huge potential to perform many accurate tests of the Standard Model (SM) with LHC data, which could yield insight into new physics mechanisms.

To make these tests, the experiments identify a particular signature  $X$  and observe, using a variety of selection criteria, a certain number of events in a given data taking period. After correcting this event rate for backgrounds and the selection efficiency, the number is converted into a cross section. The cross section,  $\sigma_{pp \rightarrow X}$  can be compared with theoretical predictions<sup>1</sup> according to the formula:  $N_{corrected} = \sigma_{pp \rightarrow X} \times L_{pp}$  where  $L_{pp}$  is the recorded proton proton luminosity.

Besides the statistical errors of a measurement, the systematic error is related to the uncertainties from the  $L_{pp}$  determination, the background and efficiency corrections within the detector acceptance and from extrapolations into the uncovered very forward rapidity regions. The interpretation of an observed cross section within the SM requires further the knowledge of the theoretical cross section. Thus the uncertainties of the proton parton distribution function (PDF) have to be considered also.

In this Section we describe the status and perspectives of the ATLAS, CMS and LHCb, the three LHC pp collision detectors [1], to determine the proton proton luminosity normalization. The investigated methods are known and studied since many years and can be separated into the absolute (1) direct and (2) indirect proton proton luminosity determination. A third approach (3) tries to measure and calculate final states only relative to well understood reactions which depend on the parton-parton luminosity and are as such largely independent of the knowledge of the pp luminosity.

- Absolute, direct or indirect, proton proton luminosity normalization: If the absolute approach is used, the interpretations of a measured reaction cross section depends still on the knowledge of parton distribution function (PDF), which must be obtained from other experiments. Examples are:

---

<sup>1</sup>Alternatively, one can also apply a Monte Carlo simulation to the theoretical prediction and compare the number of background corrected events directly.

- The proton proton luminosity normalization is based on the measurements of the beam currents and shapes. While the beam currents can be accurately determined using beam transformers, the beam profiles are more difficult to determine directly and usually constitute the dominant source of uncertainty on a luminosity measurement using this technique. The use of the machine luminosity determination using beam parameter measurements [2] and [3] will be described in Section 3.1. Alternatively one can try to measure the beam profiles also within the experiments using the precision vertex detectors. A short description of this idea, currently pursued within the LHCb collaboration, is also given in Section 3.1.
  - The simultaneous measurements of a pair of cross sections that are connected with each other quadratically via the optical theorem. A well known example of this is the measurement of the total inelastic cross section and the elastic cross section at very high pseudorapidities  $|\eta| \approx 9$  and will be described in Section 3.3. So called instantaneous or real time luminosity measurements are based on “stable” high rate measurements of particular final state reactions. Once the ratio of such reactions to the pp luminosity determination has been measured, those reactions can be subsequently used as independent luminosity monitors. Some possibilities are discussed in Section 3.4.
  - The indirect absolute proton proton luminosity normalization is based on the theoretically well understood “two photon” reaction  $pp \rightarrow pp\mu\mu$  [4, 5] (Section 3.5). This reaction could perhaps be considered as the equivalent of the luminosity counting in  $e^+e^-$  experiments using forward Bhabha scattering.
- Indirect pp luminosity measurements use final states, so called “standard candles”, with well known theoretical cross sections (Section 4). Obviously, the resulting proton proton luminosity can only be as good as the theoretical and experimental knowledge of the “standard candle” reaction. The theoretically and experimentally best understood LHC reactions are the inclusive production of W and Z bosons with subsequent leptonic decays. Their large cross section combined with experimentally well defined final states, e.g. almost background free Z and W event samples can be selected over a relative large rapidity range, makes them the preferred LHC “standard candle” reaction. Other interesting candidates are the high  $p_t$  jet - boson ( $= \gamma, W$  or  $Z$ ) final states. The indirect luminosity method requires also some knowledge of the PDFs, and of course, if one follows this approach, the cross section of the “standard candle” reaction becomes an input and can not be measured anymore. Thus, only well understood reactions should be considered as candidate reactions.
  - pp luminosity independent relative rate measurements using “standard candle” reactions. In addition to the above indirect pp luminosity determinations, “standard candle” reactions allow to perform luminosity independent relative event rate calculations and measurements. This approach has already been used successfully in the past and more details were discussed during the past HERA-LHC workshop meetings [6]. For some reactions, this approach appears to be much easier and more accurate than standard cross section measurements and their interpretations. Perhaps the best known example at hadron colliders is the measurement and its interpretation of the production ratio for Z and W events,

where Tevatron experiments have reached accuracies of about 1-2% [7,8]. Another example is related to relative branching ratio and lifetime measurements as used for b-flavored hadrons.

Furthermore the rapidity distributions of leptonic W and Z decays at the LHC are very sensitive to the PDF parameterization and, as was pointed out 10 years ago [9], one can use these reactions to determine the parton luminosity directly and very accurately over a large  $x$  (= parton momentum/proton momentum) range. In fact, W and Z production with low transverse momentum were found in this analysis to be very sensitive to  $q\bar{q}$  luminosities, and the jet-boson final states, e.g. the jet- $\gamma$ , Z, W final states at high transverse momentum are sensitive to the gluon luminosity.

In the following we attempt to describe the preparations and the status of the different luminosity measurements and their expected accuracies within ATLAS, CMS and LHCb. Obviously, all these direct and indirect methods should and will be pursued. In Section 5 we compare the advantages and disadvantages of the different methods. Even though some methods look more interesting and rewarding than others, it should be clear from the beginning that as many independent pp luminosity determinations as possible need to be performed by the experiments.

We also try to quantify the systematic accuracies which might be achieved over the next few years. As these errors depend somewhat on the overall achieved luminosity, we need in addition a hypothetical working scenario for the first 4 LHC years. We thus assume that during the first year, hopefully 2009, data at different center of mass energies can be collected by ATLAS and CMS. During the following three physics years we expect that 10 TeV will be the highest collision energy in year I and that at most  $100 \text{ pb}^{-1}$  can be collected. We assume further that during the following two years the design energy of 14 TeV can be achieved and that a luminosity of about  $1 \text{ fb}^{-1}$  and  $10 \text{ fb}^{-1}$  can be collected respectively per year. During the first few years similar numbers are expected for the LHCb experiment. However once the LHC reaches the first and second phase design luminosity of  $10^{33}/\text{cm}^2/\text{sec}$  and  $10^{34}/\text{cm}^2/\text{sec}$  it is expected that the LHCb experiment will run at an average luminosity of  $2 \times 10^{32}/\text{cm}^2/\text{sec}$  (resulting in about  $2 \text{ fb}^{-1}/\text{per year}$ ).

## 2 Luminosity relevant design of ATLAS/CMS and LHCb

In the following we give a short description of the expected performance with respect to lepton and jet identification capabilities. Especially the electron and muon measurement capabilities are important for the identification of events with leptonic decays of W and Z bosons.

Both ATLAS and CMS are large so called omni purpose experiments with a large acceptance and precision measurement capabilities for high  $p_t$  electrons, muons and photons. Currently, the simulations of both experiments show very similar performance for a large variety of LHC physics reactions with and without jets. For the purpose of this Section we focus on the possibility to identify the production of inclusive W and Z decays with subsequent decays to electrons and muons. Both experiments expect excellent trigger accuracies for isolated leptons and it is expected that electrons and muons with momenta above 20-25 GeV can be triggered with high efficiency and up to  $|\eta|$  of about 2.5. The special design of the ATLAS forward muon spectrometer should allow to detect muons with good accuracy even up to  $|\eta|$  of 2.7.

The operation of ALFA, a very far forward detector placed about 240 m down the beam line, is envisaged by the ATLAS collaboration to provide an absolute luminosity measurement, either using special optics LHC running and the use of the optical theorem or using the total cross section measurement from the dedicated TOTEM experiment installed near CMS; results from this device can be expected from 2010 and on-wards. In addition to absolute luminosity measurements from ALFA the two detectors LUCID and the Zero-Degree-Calorimeter (CDC) [10] are sensitive to the relative luminosity at time scales of single bunch crossings.

A similar approach for absolute and relative luminosity measurements is foreseen by the CMS experiment. Here it is planned that dedicated forward detectors, the Hadron Forward Calorimeter (HF) and the ZDC device provide similar results as the ones in ATLAS.

Another technique that is expected to be available early on is a luminosity-independent measurement of the  $pp$  total cross section. This will be done using a forward detector built by the TOTEM experiment [11].

The LHCb experiment [12] has been designed to search for New Physics at the LHC through precision measurements of CP violating observables and the study of rare decays in the b-quark sector. Since the  $b\bar{b}$  pairs resulting from the proton-proton collisions at the LHC will both be produced at small polar angles and in the same forward or backward cone, LHCb has been designed as a single-arm forward spectrometer covering the pseudo rapidity range  $1.9 < \eta < 4.9$ . The LHCb tracking system, which is composed of a silicon vertex detector, a warm dipole magnet and four planar tracking stations, will provide a momentum resolution of  $\delta P/P = (0.3 + 0.0014P/GeV)\%$  [13]. Muon identification is primarily achieved using a set of five planar multi-wire proportional chambers, one placed in front of the calorimeter system and four behind, and it is expected that for the momenta range 3-150GeV/c an identification efficiency of  $\sim 98\%$  and an associated pion dis-identification rate of  $\sim 1\%$  will be achieved. The reconstruction of primary and secondary vertices, a task of crucial importance at b physics experiments, will be virtually impossible in the high particle multiplicity environment present with the nominal LHC running luminosity of  $10^{34} cm^{-2} s^{-1}$  - LHCb has therefore been designed to run at the lower luminosity of  $2 \times 10^{32} cm^{-2} s^{-1}$ .

Recent LHCb simulations have shown that leptonic W and Z decays to muons can be identified with a small background in the forward and very forward rapidity region starting from  $\eta$  of 1.9 and up to values larger than 4. As will be discussed later in more detail, the common muon acceptance region for the three LHC experiments between 1.9 and about 2.5 will allow to cross check and normalize the W and Z measurements in this region. Consequently the unique large rapidity from 2.5 to 4.9 can be used by LHCb to investigate the very low x range of the PDFs for the first time.

The absolute luminosity at LHCb will be obtained either directly, by making measurements of the beam parameters, or indirectly via a measurement of the event rate of an accurately predicted physics process.

As will be explained in the following Sections, all experiments will try to perform as many as possible direct and indirect absolute and relative luminosity measurements and will, if available, at least during the first years, also use luminosity numbers from the machine group.

## 2.1 Lepton triggering and W/Z identification.

Generally, the lepton trigger selections depend on the instantaneous luminosity and some pre-scaling might eventually be needed. However, current simulations by all experiments show that the envisaged  $|\eta|$  and  $p_t$  thresholds will not limit the measurement accuracies of leptons originating from W and Z decays.

The lepton trigger selections that are generally perceived to be used for most W and Z related analysis are very similar in ATLAS and CMS as indicated in Table 1.

Experiment	Trigger selection $e$		Trigger selection $\mu$	
	$p_T$	$ \eta $	$p_T$	$ \eta $
ATLAS	25 GeV	2.5	20 GeV	2.7
CMS	20 GeV	2.5	20 GeV	2.1
LHCb*	–	–	2.5 GeV	1.9-4.9

Table 1: For ATLAS and CMS the lepton trigger/selection  $p_t$  thresholds are given for single isolated leptons. \*For the LHCb threshold is given for the muon pair mass instead of single muons and only positive values of  $\eta$  are covered.

Trigger and reconstruction efficiencies for leptonic W and Z decays within the acceptance of the detectors have been estimated for ATLAS to be 97.7% and 80.0% for electrons and 84.3% and 95.1% for muons, respectively. The reconstruction efficiency includes the trigger efficiencies and the off-line electron and muon selections used later to identify clean inclusive W and Z event samples [14].

The current equivalent trigger and off-line efficiencies for CMS are about 85% and 77% for electrons and combined about 85% for single muons [15]. Similar efficiency numbers for muons from W and Z decays are expected within the LHCb acceptance region [16]. Current simulations show that these numbers can be determined with high accuracies, reaching perhaps 1% or better, at least for isolated leptons<sup>2</sup> which have a transverse momentum some GeV above the trigger thresholds. For lower momenta near the thresholds or for additional special trigger conditions somewhat larger systematic uncertainties can be expected.

## 3 Direct and indirect absolute pp luminosity measurements

Three different absolute proton proton luminosity measurements are discussed in this Section. (1) The machine luminosity determination using beam parameter measurements [17], (2) the luminosity independent total pp cross section measurement combined with the measurement of the elastic pp scattering rate [11] and (3) the measurement of the “two photon” reaction  $pp \rightarrow pp\mu\mu$  [4, 5]. As will be discussed in more detail in Section 5, only method (3) can be performed during the normal collision data taking. For method (1) some special methods, which take the actual detector performance during each run into account, need to be developed. Method 2 uses a two phase approach (a) a special machine optics run with low luminosity to determine the total

<sup>2</sup>As isolated high  $p_t$  photons are triggered essentially like electrons similar accuracies for both particle types can be assumed.

### 3.1 Proton-proton luminosity from machine parameters<sup>3</sup>

The luminosity for colliding beams can be directly obtained from geometry and numbers of particles flowing per time unit [2]. This can be used to determine the absolute LHC luminosity from machine parameters without prior knowledge of pp scattering cross sections. The principle is briefly outlined here. More details can be found in [3].

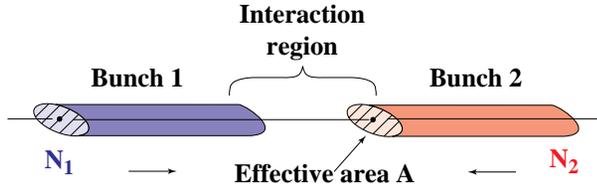


Fig. 1: Luminosity from particles flux and geometry.

For two bunches of  $N_1$  and  $N_2$  particles colliding head-on in an interaction region as sketched in Fig.1 with the frequency  $f$  the luminosity is given as

$$\mathcal{L} = \frac{N_1 N_2 f}{A_{\text{eff}}} . \quad (1)$$

$A_{\text{eff}}$  is the *effective transverse area* in which the collisions take place. For a uniform transverse particle distribution,  $A_{\text{eff}}$  would be directly equal to the transverse beam cross section. More generally, the effective area can be calculated from the overlap integral of the two transverse beam distributions  $g_1(x, y)$ ,  $g_2(x, y)$  according to

$$\frac{1}{A_{\text{eff}}} = \int g_1(x, y) g_2(x, y) dx dy . \quad (2)$$

For equal Gaussian beams

$$g_1 = g_2 = \frac{1}{2\pi\sigma_x\sigma_y} \exp \left[ -\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2} \right] \quad (3)$$

we obtain for head-on collisions  $A_{\text{eff}} = 4\pi\sigma_x\sigma_y$  so that

$$\mathcal{L} = \frac{N_1 N_2 f}{4\pi\sigma_x\sigma_y} . \quad (4)$$

The collision frequency  $f$  is accurately known. The number of particles circulating in a storage ring is measured using beam current transformers to roughly 1% precision [17].

The main uncertainty in the absolute luminosity determination from machine parameters is expected to originate in the knowledge of the transverse beam dimensions. Safe operation

---

<sup>3</sup>Contributing author: H. Burkhardt

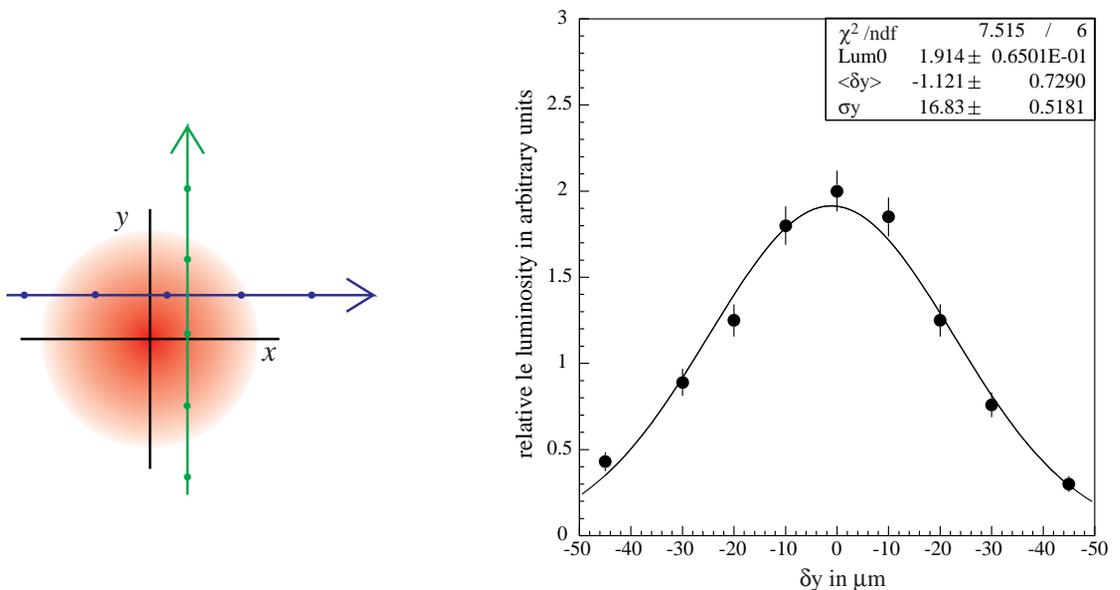


Fig. 2: Schematic view of the steps involved in an orthogonal separation scan proposed for the LHC (left) and a possible result in one direction (based on early LEP data) shown on the right.

of the LHC requires a rather good knowledge of the optics and beam sizes and we expect that this should already allow a determination of the luminosity from machine parameters to about 20 – 30 percent. A much better accuracy can be obtained when the size of the overlap region at the interaction points is determined by measuring the relative luminosity as a function of lateral beam separation, as illustrated in Fig. 2. This technique was pioneered at the ISR [18] and allowed to reduce the uncertainty to below 1%, [19, 20].

For the more complicated LHC and early operation, a 10% overall uncertainty in the absolute LHC machine luminosity calibration should be a realistic goal. The actual precision will depend on the running time and effort which is invested. A relatively small number of scans under favorable beam conditions will in principle be sufficient to obtain and verify the reproducibility in the absolute luminosity calibration. While fast scans may always be useful to optimize collisions, we assume that any dedicated, detailed luminosity scans will become obsolete when the other, cross section based luminosity determinations described in these proceedings allow for smaller uncertainties.

Optimal running conditions are moderate bunch intensities, large bunch spacings, no crossing angle and  $\beta^* = 2$  m or larger. These conditions are in fact what is proposed anyway for the initial LHC operation with 43 – 156 bunches per beam. Statistics are not expected to be a problem. For early operation at top energy (10 - 14 TeV) with 43 bunches and  $4 \times 10^{10}$  particles per bunch, before beams are squeezed. at a  $\beta^* = 11$  m, we already expect luminosities of the order of  $10^{30} \text{ cm}^{-2} \text{ s}^{-1}$  resulting in event rates of  $10^4$  Hz, for a cross section of 0.01 barn as typical for the low angle luminosity monitors.

From the LHC injectors, we expect bunch by bunch variations of about 10% in intensity and 20% in emittance. For the large spacing between bunches in the operation with up to 156

bunches, there is no need for crossing angles at the interaction points. Parasitic beam-beam effects will be negligible. All bunches in each beam will follow the same equilibrium orbit and collide at the same central position.

Calibration runs require good running conditions and in particular good beam lifetimes. Bunch by bunch differences are not expected to change significantly during a scan. Storing bunch intensities at the beginning and end of a scan and using one set of timed averaged bunch intensities for a scan should be sufficient. To avoid any bias, it will be important to use the correct pairing of bunch intensities and relative luminosities in the calculation of absolute bunch luminosities according to Eq. 1, before any summing or averaging over different bunches.

We are currently preparing an on-line application for automatic luminosity scans<sup>4</sup>. Scan parameters like range, step size and duration can be set before the start of the scan. Once the parameters are defined, it is possible to launch automatic horizontal and vertical separation scans in the LHC interactions regions. For a detailed scan, we may choose a range from  $-4$  to  $+4 \sigma$  in nominal beam size in steps of  $0.5 \sigma$ , resulting in 17 equidistant points. If we wait 1 s between points to allow for the magnets to change and for 2 s integration time, the scan time would still be below a minute per plane. Details are currently being worked out in close collaboration with the experiments. Exchanging all data bunch-by-bunch at a 1 Hz rate between the machine control room (CCC) and the experiments would be rather demanding and risks to saturate current capacities.

For the initial running, it will be sufficient to exchange average values at about 1 Hz rate. It allows quality monitoring and the determination of the peak position. For the detailed off line analysis, we only have to rely on local logging and timing information synchronized to at least 1 s precision at the beginning of the scan. With fixed time interval defined and saved before the scan, this allows for off-line synchronization of the detailed data and a complete bunch by bunch analysis.

### 3.2 Direct measurements of the absolute luminosity at LHCb

LHCb plans to measure the absolute luminosity using both the Van Der Meer scan, [18], and beam-gas techniques following a more recently proposed method [21]. Here one tries to determine the transverse beam profiles at colliding beam experiments utilizing the precision vertex detectors found at modern HEP experiments to reconstruct beam gas interactions near the beams crossing point. The vertex resolution in the transverse direction at LHCb can be parameterized by the relation

$$\sigma_{x,y} = \frac{100 \mu m}{\sqrt{N_{tracks}}} \quad (5)$$

where  $N_{tracks}$  is the number of tracks originating from the vertex. Since the nominal transverse bunch size at LHCb will be  $100 \mu m$ , the reconstruction of beam-gas vertices's, which will have a track multiplicity of  $\sim 10$ , will enable the measurement of the colliding bunch profiles and the beam overlap integral. This method is currently under investigation by the LHCb collaboration and is expected to result in a luminosity measurement with an associated uncertainty of 3-5%.

---

<sup>4</sup>Done by Simon White, as part of his PhD thesis work on the LHC machine luminosity determination

### 3.3 Absolute pp luminosity from specialized detectors and from the total cross section measurement

ATLAS and CMS are planning to perform absolute and relative pp luminosity measurements using dedicated luminosity instruments.

Three particular luminosity instruments will operate around the ATLAS interaction point. The absolute luminosity measurement will be provided by ALFA [10] placed 240m down the beam line and due to operate in 2010. This measurement requires some special optics low luminosity running of the LHC and should be able to measure the very low angle Coulomb scattering reaction. The expected precision is of the order 3%, depending on yet unknown LHC parameters during running. The ALFA detector can also measure the absolute luminosity using the optical theorem if the Coulomb region can not be reached. Extrapolating the elastic cross section to very low momentum transfer  $t = 0$  and using the total cross section as measured by TOTEM [11] (located at the CMS interaction point) current simulations indicate that a precision of about 3% might also be reached with this method. In addition to absolute luminosity measurements from ALFA, LUCID and a Zero-Degree-Calorimeter (ZDC) [10] are sensitive to the relative single bunch crossings luminosity. LUCID and ZDC will however not give absolute measurements.

A similar approach is currently foreseen by the CMS collaboration [22].

### 3.4 Real time relative luminosity measurements

A large number of instantaneous relative luminosity measurements have been discussed during the past years by ATLAS, CMS and LHCb and more details can be found in the three presentations given during the “standard candle” session of this workshop [23]. As an example we outline in the following some ideas discussed within CMS.

Multiple techniques capable of providing suitable luminosity information in real time have been identified in CMS. One technique employs signals from the forward hadron calorimeter (HF) while another, called the Pixel Luminosity Telescope (PLT), uses a set of purpose-built particle tracking telescopes based on single-crystal diamond pixel detectors. At this writing, the PLT has not been formally approved, but is under study. The methods based on signals from the HF described are the ones being most vigorously pursued.

Two methods for extracting a real-time relative instantaneous luminosity with the HF have been studied. The first method is based on “zero counting,” in which the average fraction of empty towers is used to infer the mean number of interactions per bunch crossing. The second method called “EtSum method” exploits the linear relationship between the average transverse energy per tower and the luminosity.

Outputs of the QIE chips used to digitize the signals from the HF PMTs on a bunch-by-bunch basis are routed to a set of 36 HCAL Trigger and Readout (HTR) boards, each of which services 24 HF physical channels. In order to derive a luminosity signal from the HTR, an additional mezzanine board called the HF luminosity transmitter (HLX) is mounted on each of the HTR boards. The HLX collects channel occupancy and  $E_T$  sum data to create eight histograms: two sets of three occupancy histograms, one  $E_T$ -sum histogram, and one additional occupancy histogram. These histograms comprise about 70 KB of data, which is transmitted at a rate of approximately 1.6 Mbps to a dedicated luminosity server via an Ethernet switch that

aggregates the data from multiple HLX boards for further processing.

Although all HF channels can be read by the HLX, MC studies indicate that the best linearity is obtained using only the inner four  $\eta$  rings. The algorithm has been optimized to minimize sensitivity to pedestal drifts, gain changes and other related effects. Both “Zero Counting” and the “EtSum” method have demonstrated linearity up to LHC design luminosity. A statistical error of about 1% will be achieved at  $\text{fewtimes} \times 10^{31} \text{cm}^{-2} \text{s}^{-1}$ . Hence the dominant error on the absolute luminosity will result from the normalization of the online relative luminosity.

### 3.5 Proton-proton luminosity from the reaction $pp \rightarrow pp\mu\mu$

The QED process  $pp \rightarrow pp\mu^+\mu^-$ , where a  $\mu^+\mu^-$  pair is produced via photon-photon scattering, was first proposed for luminosity measurements at hadron colliders in [4]. At the LHC such pairs will be predominantly produced with small transverse momenta, at small polar angles and in the same forward or backward cone.

All three experiments are considering to use the well calculated  $pp \rightarrow pp\mu\mu$  process for measuring absolute luminosity. The theoretical understanding of this QED photon-photon scattering reactions is considered to be accurate to better than 1%. Consequently this final state is thus often considered to be the perfect theoretical luminosity process. However, the experimental identification of this process requires to select muon pairs with low mass and within a well understood acceptance. The measurement of this reaction at a hadron collider appears to be much more difficult than the corresponding measurements of the reaction  $ee \rightarrow ee\mu\mu$  at LEP. The systematic measurement error for example in L3 and after several years of data taking was about  $\pm 3\%$  [24]

Current simulations by the three LHC experiments indicate that the final state can be identified using straight forward criteria. For ATLAS and CMS one finds that about 1000 accepted events could at best be expected for an integrated luminosity of  $1 \text{ fb}^{-1}$ , resulting in a statistical error of about  $\pm 3\%$ .

For example the ATLAS study selects oppositely charged back-to-back muon tracks with  $p_T > 6 \text{ GeV}$  and  $|\eta| < 2.2$  with an invariant mass less than  $60 \text{ GeV}$  and a common vertex with no other tracks originating from it (isolation), yields a cross section of  $1.33 \text{ pb}$ . Thus, about 1300 events can be expected for running periods with a luminosity of  $1 \text{ fb}^{-1}$  and yielding a potential statistical error of 3%. However, backgrounds not only from pile up events will be a critical issue. Some proton tagging with high luminosity roman pots is currently investigated but this will certainly reduce the accepted cross section and introduce additional acceptance errors. Similar conclusions have been reached by simulations performed within the CMS collaboration. Consequently, both experiments expect that, during the coming years, this reaction will be mainly used as a cross check of the other methods.

The cross section for this process where both muons lie inside the LHCb acceptance and have a combined invariant mass greater than  $2.5 \text{ GeV}$  is  $\approx 88 \text{ pb}$ . The expected uncertainty is perhaps 1% or smaller and comes mainly from rescattering corrections [5], i.e. strong interactions between the interacting protons.

The feasibility of using the elastic two photon process  $pp \rightarrow p + \mu^+\mu^- + p$  to make luminosity measurements at LHCb was first explored in [25] and has recently been investigated in

more detail by members of the LHCb collaboration [26]. A variety of background processes have been studied: dimuons produced via inelastic two-photon fusion and double pomeron exchange; low mass Drell-Yan pairs; QCD processes such as  $b\bar{b} \rightarrow \mu^+\mu^- + X$ ; and the combinatoric backgrounds caused by  $K/\pi$  mis-identification. A simple offline selection has been developed that requires: the dimuon pair transverse momentum to be less than 50MeV/c; the dimuon invariant mass to be in the range  $2.5\text{GeV}/c^2 < M_{\mu\mu} < 20\text{GeV}/c^2$ ; and a charged particle multiplicity of less than 3 (i.e. the event should contain a  $\mu^+\mu^-$  pair and no other charged particles). These criteria select  $\sim 27\%$  of the signal events that pass the trigger and are reconstructed and result in a background contamination that is  $(4.1 \pm 0.5(\text{stat.}) \pm 1.0(\text{syst.}))\%$  of the signal level with the dominant contribution due  $K/\pi$  mis-identification. Overall it is expected that  $\sim 10^4$   $pp \rightarrow p + \mu^+\mu^- + p$  events will be triggered, reconstructed and selected at LHCb during one nominal year of data taking ( $2\text{fb}^{-1}$ ). Systematic uncertainties on a luminosity measurement at LHCb using this channel are estimated to be  $\sim 1.31\%$  and are dominated by the uncertainty on the predicted cross section for events containing dimuons produced via double pomeron exchange, an uncertainty that is expected to be reduced in the near future. A measurement of the absolute luminosity at LHCb using this channel and a dataset of  $2\text{fb}^{-1}$  will therefore be possible with an associated uncertainty of  $\sim 1.5\%$ .

In summary, the accurate measurement of this theoretically well understood reaction looks like an interesting challenge for the LHC experiments. Interesting results can be expected once integrated luminosities of  $5\text{fb}^{-1}$  and more can be accumulated for ATLAS and CMS and about  $1\text{fb}^{-1}$  for LHCb. Of course, it remains to be proven, if the systematic uncertainties under real data taking conditions can indeed be reduced to the interesting 1% level.

#### 4 Indirect and relative pp luminosity measurements

The methods to measure the absolute proton proton luminosity and their limitations have been described in the previous chapter.

In this Section we will describe the possibilities to measure the luminosity indirectly using well defined processes, so called ‘‘Standard Candles’’ and their use to further constrain the PDFs and discuss the possibility to ‘‘measure’’ directly the parton-parton luminosities.

Before describing the details of these indirect approaches, a qualitative comparison of luminosity measurements at  $e^+e^-$  colliders and hadron colliders might be useful. The most important difference appears to be that in the  $e^+e^-$  case one studies point like parton parton interactions. In contrast, at hadron hadron interactions one studies the collision of protons and other hadrons made of quarks and gluons. As a result, in one case the Bhabha elastic scattering reaction  $e^+e^- \rightarrow e^+e^-$  at low  $Q^2$  reaction can be calculated to high accuracy and the observed rate can be used as a luminosity normalization tool. In contrast, the elastic proton proton scattering cross section can not be calculated at the LHC nor at any other hadron colliders. As a consequence, absolute normalization procedures depend always on the measurement accuracy of the pp total cross section. Even though it is in principle possible to determine the pp total cross section in a luminosity independent way using special forward detectors like planned by the TOTEM or the ALFA experiments, the accuracy will be limited ultimately and after a few years of LHC operation to perhaps a few %.

Furthermore, as essentially all interesting high  $Q^2$  LHC reactions are parton parton collisions, the majority of experimental results and their interpretation require the knowledge of parton distribution functions and thus the parton luminosities.

Following this reasoning, more than 10 years ago, the inclusive production of W and Z bosons with subsequent leptonic decays has been proposed as the ultimate precision parton parton luminosity monitor at the LHC [9]. The following points summarize the arguments why W and Z production are indeed the ideal “Standard Candles” at the LHC.

- The electroweak couplings of W and Z bosons to quarks and leptons are known from the LEP measurements to accuracies smaller than 1% and the large cross section of leptonic decays W and Z bosons allows that these final states can be identified over a large rapidity range with large essentially background free samples.
- Systematic, efficiency corrected counting accuracies within the detector acceptance of 1% or better might be envisioned during the early LHC running. In fact it is believed that the relative production rate of W and Z can be measured within the detector acceptance with accuracies well below 1%.
- Theoretical calculations for the W and Z resonance production are the most advanced and accurately known LHC processes. Other potentially more interesting LHC reactions, like various diboson pair production final states are expected to have always larger, either statistical or systematic, experimental and theoretical uncertainties than the W and Z production.
- The current PDF accuracies, using the latest results from HERA and other experiments demonstrate that the knowledge of the quark and anti quark accuracies are already allowing to predict the W and Z cross at 14 TeV center of mass energies to perhaps 5% or better. The measurable rapidity and  $p_t$  distributions of the Z boson and the corresponding ones for the charged leptons from W decays can be used to improve the corresponding parton luminosity functions.

Obviously, the use of W and Z bosons as a luminosity tool requires that the absolute cross section becomes an input, thus it can not be measured anymore. As a result this method has been criticized as being “a quick hack at best”. In contrast, advocates of this method point out that this would not be a noticeable loss for the LHC physics program.

#### 4.1 Using the reaction $pp \rightarrow Z \rightarrow \ell^+ \ell^-$ to measure $L_{pp}$

Very similar and straight forward selection criteria for the identification of leptonic Z decays, depending somewhat on the detector details and the acceptance region, are applied by ATLAS, CMS and LHCb. In the following the current selection strategy in ATLAS and LHCb are described.

#### 4.2 Measuring Z and W production, experimental approaches in ATLAS

The ATLAS W and Z cross section measurements are based on the following selections in the electron and muon channels:

- A typical selection of  $W \rightarrow e\nu$  requires that events with “good” electrons have to fulfill the additional kinematic acceptance criteria:

$p_T > 25$  GeV,  $|\eta| < 1.37$  or  $1.52 < |\eta| < 2.4$ .

The criteria for  $W \rightarrow \mu\nu$  muons are similar where  $p_T > 25$  GeV and  $|\eta| < 2.5$  is required. Furthermore, in order to classify the event as a  $W$  event, the reconstructed missing transverse momentum and the transverse mass should fulfill  $E_T(miss) > 25$  GeV and  $m_T(W) > 40$  GeV.

- The selection of  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  requires that a pair of oppositely charged electrons or muons is found. Due to lower background the electrons should have  $p_T > 15$  GeV and  $|\eta| < 2.4$  and their invariant mass should be between 80-100 GeV.

Similar criteria are applied for the muons with  $p_T > 15$  GeV and  $|\eta| < 2.5$ . The reconstructed mass should be between 71-111 GeV.

Following this selection and some standard Monte Carlo simulations, the expected number of reconstructed events per  $10 \text{ pb}^{-1}$  at  $\sqrt{s} = 14$  TeV are about 45000, 5500 for W and Z decays to electrons and 60000, and 5000 for the decays to muons, respectively. Thus, even with a small data sample of only  $10 \text{ pb}^{-1}$ , the statistical uncertainty for the Z counting comes close to 1% in each channel.

Systematic uncertainties from the experimental selection are dominated by the Z efficiency determination and from backgrounds in the W selection. Other sources of uncertainties originate from the knowledge of energy scale and the resolution. The lepton efficiencies are evaluated by considering  $Z \rightarrow \ell\ell$  events and using the so called ‘‘tag and probe’’ method, like for example described by the D0 experiment [7, 8]. The efficiency uncertainty associated with the precision of this method has been estimated for a data sample of  $50 \text{ pb}^{-1}$  ( $1 \text{ fb}^{-1}$ ) of data to be 2% (0.4%) for W and 3% (0.7%) for Z events. The backgrounds for W events are of the order 4% in the electron channel and 7% in the muon channel. The main contributions are from other W or Z decays, and are thus well understood, leading to background uncertainties of the order 4% for both channels if a sample  $50 \text{ pb}^{-1}$  is analyzed. For much larger samples it is expected that uncertainties at or below 1% can be achieved. The backgrounds for the Z decays are very small, and can be determined accurately from mass spectrum, and hence does not carry any sizable uncertainty. It has been demonstrated, that the detector scales and resolutions can be determined very accurately [14], and the associated uncertainties are therefore also close to negligible. Some detailed studies demonstrate that eventually the systematic error between 1-2% or even smaller might be achieved for the W and Z counting and within the detector acceptance up to rapidities of about 2.5.

In order to use this number for the pp luminosity determination the total inclusive W and Z cross-section at NNLO can be used. These have been calculated to be 20510 pb and 2015pb, respectively [27]. Variations in models, floating parameters, and other theoretical uncertainties lead to significant variations in the estimates. The uncertainties on these calculation are estimated to be 5% or smaller. This uncertainty appears to be currently dominated by the PDF uncertainties needed to extrapolate to the experimentally uncovered large rapidity region. More discussions about these uncertainties can be found for example at [28] and [29].

It can be assumed that the detailed studies of the rapidity distributions within the acceptance region with W and Z decays might eventually lead to further error reductions.

### 4.3 Measuring Z production, experimental approach in LHCb

The uncertainty on the predicted Z production cross section at the LHC comes from two sources: the uncertainty on the NNLO partonic cross section prediction [27], which contributes an uncertainty of  $< 1\%$ , and uncertainties in our understanding of the proton Parton Distribution Functions (PDFs) which, for the latest MSTW fit [30], contribute an uncertainty of  $\sim 3\%$  for Z bosons produced with rapidities in the range  $-5 < y < 5$ .

A measurement of the Z production rate at LHCb via the channel  $Z \rightarrow \mu^+\mu^-$ , which provides a final state that is both clean and fully reconstructible, can be achieved with high efficiency and little background contamination. In addition, since the dimuon trigger stream at LHCb [31] requires two muons with an invariant mass larger than 2.5 GeV and a summed transverse momentum ( $P_T^1 + P_T^2$ ) greater than 1.4 GeV, a high trigger efficiency of  $\sim 95\%$  is expected for these events. A variety of background sources for this channel have been investigated: other electroweak processes such as  $Z \rightarrow \tau^+\tau^-$  where both taus decay to muons and neutrinos; QCD processes such as  $b\bar{b} \rightarrow \mu^+\mu^- + X$ ; and events where two hadrons with an invariant mass near the Z mass are both mis-identified as muons. To deal with these backgrounds an off-line selection has been developed [32] that requires: the dimuon invariant mass to be within 20 GeV of the Z mass; the higher and lower transverse momentum muons to be greater than 20 GeV and 15 GeV respectively; the impact parameter of both muons is consistent with the primary vertex; and both muons have associated hadronic energy that is less than 50 GeV. For  $Z \rightarrow \mu^+\mu^-$  events that are triggered and reconstructed at LHCb, these off-line selection criteria will select  $91 \pm 1\%$  of the signal events while reducing the background to  $(3.0 \pm 2.9)\%$  of the signal level with the dominant contribution due to the combinatoric backgrounds from pion and kaon mis-identification. It is expected that these backgrounds can be well understood from real data or removed using muon isolation criteria. Overall it is expected that  $Z \rightarrow \mu^+\mu^-$  events will be triggered, reconstructed and selected at LHCb at a rate of  $\sim 190 \text{ evts}/\text{pb}^{-1}$ . Systematic uncertainties have also been investigated and it is expected that with as little as  $5 \text{ pb}^{-1}$  of data the experimental efficiency (trigger, tracking, muon identification etc.) can be measured with an uncertainty of  $\sim 1.5\%$  enabling a luminosity measurement with an uncertainty of  $\sim 3.5\%$ .

### 4.4 PDF and relative parton-parton luminosity measurements

Theoretically well understood reactions at the LHC offer the possibility to use their rapidity distributions to improve today's knowledge of PDFs. Especially the resonance production of W and Z bosons with leptonic decays with low and high transverse momentum and the production of isolated high  $p_t$   $\gamma$ -Jet events have been demonstrated to be very sensitive to the relative parton distribution functions. Simulations from ATLAS and CMS have shown that experimental errors on these rapidity regions up to  $|y|$  of about 2.5 can probably be performed with accuracies eventually reaching perhaps 1% or better. The possibility to cross-check the measurements with W and Z decays to (a) electron(s) and (b) muon(s) and between both experiments will of course help to reach the accuracy.

During the past years simulation studies from the LHCb collaboration have shown that the experiment has a unique potential to extend the acceptance region from ATLAS and CMS for muons up to rapidity values at least up to 4.5. Furthermore, the existing overlap region for y

between 1.9 and 2.5 should allow to reduce normalisation uncertainties. Obviously, these rapidity values are understood as being reasonably accurate but qualitative values and more precise values will be defined once real data will allow to define a well understood fiducial volume of the detectors.

In addition, the LHCb collaboration has investigated the possibility to identify clean samples of very low mass Drell-Yan mu-pair events. The results indicate that such pairs can be measured within their acceptance region down to masses of 5 GeV. Such a measurement would in principle allow to measure PDFs for  $x$  values approaching extremely low values of  $10^{-6}$  for the first time [33].

It should be clear that such measurements, which are known to be very sensitive to quark, antiquark and gluon relative parton luminosities will not allow an absolute PDF normalisation. Such an improvement of absolute PDF normalisation would require the accurate knowledge of the proton-proton luminosity to better than today's perhaps  $\pm 3\%$  PDF accuracy obtained from the HERA measurements over a large  $x$  range and obviously lower  $Q^2$ . The alternative approach to combine the relative parton luminosities over the larger  $x, Q^2$  range using the sum rules has, to our knowledge, so far not been studied in sufficient detail.

A more detailed analysis of the different experimental approaches to improve the PDFs are interesting but are beyond the scope of this note about the luminosity. Nevertheless we hope that the experimentalists of the three collaborations will start to combine their efforts and will pursue the PDF measurements, in direct collaboration with theorists, during the coming years.

## 5 Comparing the different pp luminosity measurements

A relatively large number of pp luminosity measurements has been proposed and the most relevant have been discussed in this note. Here we try to give a critical overview of the different methods and their potential problems. Despite these advantages and disadvantage it should be clear that it is important to perform as many as possible independent luminosity methods during the coming years.

- **The machine luminosity determination using beam parameters:**

This method will be pursued independently of the experiments and its main purpose will be to optimize the performance of the LHC and thus providing a maximum number of physics collisions for the experiments. The potential to use this number as an almost instantaneous absolute luminosity number with uncertainties of perhaps  $\pm 10\%$  (and eventually  $\pm 5\%$ ), assuming that non gaussian tails of the beam can be controlled to this accuracy will certainly be useful to the experiments. Of course the experiments would lose somewhat their "independence" and still need to combine this number with their actual active running time.

However, one should remember that the Tevatron experiments did not use this method for their measurements.

The method to determine the beam size using the LHCb precision vertex detector look very promising and it is hoped that their approach might result in a pp luminosity measurement with an associated uncertainty of 3-5%.

- **Total cross section and absolute luminosity normalisation with specialized far for-**

**ward Detectors:**

The luminosity independent total pp cross section measurement is planned by the TOTEM collaboration and by the ALFA detector. Using these numbers both ATLAS and CMS plan to obtain the pp luminosity from the counting of the pp elastic scattering counting numbers from the forward detectors which thus depend on the knowledge of the total cross section measurement. In order to obtain this number some few weeks of special optics and low luminosity LHC running are required. As all LHC experiments are very keen to obtain as quickly as possible some reasonable luminosity at 14 TeV center of mass energy it is not likely that those special LHC data taking will happen during the first year(s) of data taking. Furthermore, despite the hope that the total cross section can be determined in principle with an interesting accuracy of  $\pm 1\%$ , it remains to be demonstrated with real LHC running. In this respect it is worth remembering that the two independent measurements of the total cross section at the Tevatron differed by 12% while much smaller errors were obtained by the individual experiments. As a result the average value with an error of  $\pm 6\%$  was used for the luminosity normalisation.

- **Luminosity determination using  $Z \rightarrow \ell\ell$ :**

This method provides an accurate large statistic relative luminosity number. It will be as accurate as the theoretical cross section calculation, which is based on the absolute knowledge of the PDFs from other experiments, from unknown higher order corrections and their incomplete Monte Carlo implementation. Today's uncertainties are estimated to be about 5%. It has been estimated, assuming the experiments perform as expected, that the potential Z counting accuracy within the acceptance region including efficiency corrections might quickly reach  $\pm 1\%$ . The extrapolation to the uncovered rapidity space, mainly due to the worse knowledge of the PDFs in this region, increases the error to perhaps 3%. Taking other theoretical uncertainties into account an error of  $\pm 5\%$  is currently estimated. Of course, advocates of the Z normalisation method like to point out that the real power of this method starts once relative measurements, covering similar partons and similar ranges of the parton distribution functions will be performed with statistical errors below 5%. Examples where such a normalization procedure looks especially interesting are the relative cross section measurements of  $N(Z)/N(W)$ ,  $N(W^+)/N(W^-)$ , high mass Drell-Yan events with respect to Z events and diboson final states decaying to leptons. Of course, correlations and anticorrelations between quark and gluon dominated production rates exist and need to be carefully investigated before similar advantages for the gluon PDFs can eventually be exploited. The loss of an independent Z cross section measurement would of course be a fact of life.

- **pp luminosity from the reaction  $pp \rightarrow pp\mu\mu$ :**

A measurement of this reaction offers in principle a direct and theoretically accurate proton proton luminosity value. Unfortunately current simulations from the experiments indicate that the accepted cross section is relatively small and only a few 1000 events can be expected per  $\text{fb}^{-1}$ . The different simulation results indicate that the backgrounds can be suppressed sufficiently without increasing the experimental systematics too much. Simulation studies [34] in CMS find that in the absence of pile-up, of the order 7000 events/fb can be selected. Apart from pile-up a leading source of systematic error is the contamina-

tion of the signal with events in which one of the protons dissociates. In the absence of pile-up, the use of the Zero-Degree-Calorimeters (one on each side of IP) and the Castor calorimeter (in 2009/10 available only on one side of the IP) in veto can improve the signal to background ratio from  $\sim 1$  to  $\sim 3$ . Hence in CMS this method may provide a means of measuring the absolute luminosity in the first LHC data with a total error of below 10%. In addition, the current simulation results indicate that small systematic errors of perhaps 1-2% might eventually be achievable<sup>5</sup> once a yearly luminosity of 5-10 fb<sup>-1</sup> in ATLAS and CMS (2 fb<sup>-1</sup> for LHCb) might be recorded. It remains to be seen if muons with transverse momenta well below 20 GeV can indeed be measured as accurately as muons with transverse momenta above 25 GeV.

### 5.1 Which luminosity accuracy might be achievable and when

Of course the potential time dependent accuracy of the different luminosity methods can only be guessed today as such numbers depend obviously on the LHC machine performance during the coming years. For the purpose of this Section we are mainly interested in measurements at the 14 TeV center of mass energy and assume that the following “data samples” would define such “years”. Of course, it could be hoped that the luminosity and energy increase would go much faster resulting in “some” shorter LHC years. Thus we assume that the first 14 TeV year, currently expected to be 2010, will correspond to 0.1 fb<sup>-1</sup>, followed by a 1 fb<sup>-1</sup> year. During the third and fourth year ATLAS and CMS expect to collect about 5 fb<sup>-1</sup> and 10 fb<sup>-1</sup> while LHCb expects to collect roughly 2 fb<sup>-1</sup> per year. We assume further that the special optics low luminosity data taking periods requiring perhaps a few weeks for TOTEM and similar for ALFA will take only place during the year when more than 1 fb<sup>-1</sup> per year or more can be expected.

As a result, for the first two 14 TeV running years, realistic luminosity numbers could come from (1) the machine group and (2) from the indirect method using the inclusive production of  $Z$  events with leptonic decays.

As has been pointed out in Section 3.1 the method (1) would, without any additional efforts by the machine group, allow a first estimate with a  $\pm 20$ -30% luminosity accuracy. We assume however that, due to the delay of the real 14 TeV start to 2010, enough resources could be found that people within the machine group could carefully prepare for the necessary beam parameter measurements and that the experiments will do the corresponding efforts to correct such a machine luminosity number for real detector data taking one could hope for a 10% measurement for 2010 and a 5% accuracy for 2011.

In contrast, method (2) would by definition be an integrated part of any imaginable experimental LHC data taking period. In fact, if enough attention is put into the  $Z$  counting method, the data expected during 2010 running might already reach statistical errors of  $\pm 2\%$  per 5 pb<sup>-1</sup> periods. Thus perhaps about 10-20 such periods could be defined during the entire year and systematic errors for the lepton efficiency correction within the detector acceptance could reach similar  $\pm 2$ -3% accuracies. During the following years these errors might decrease further to 1% or better. Once the rate of any “stable” simple high rate final states and even trigger rates relative

---

<sup>5</sup>It might be interesting to study the experience from similar measurements at the experimentally ideal conditions of LEP, where uncertainties above  $\pm 3\%$  have been reported [24].

to the  $Z$  counting rate has been determined, such relative event rates can be used subsequently to track the “run” luminosity and even the real time luminosity with similar accuracy.

Theoretical limitations of the cross section knowledge, not expected to improve without LHC data taking, would limit the accuracy to about  $\pm 5\%$ . The expected detailed analysis of the 2010 rapidity distributions of  $W$ ,  $Z$  and  $\gamma$ -jet events will allow some improvements for the years 2011 and beyond. We can thus expect that appropriate ratio measurements like the cross section ratio measurements of  $Z/W^\pm$  and  $W^-/W^+$  will already reach systematic accuracies of  $\pm 1\text{-}2\%$  during 2010 and 1% or better in the following years. Measurement of  $b$  physics, either in LHCb or in ATLAS and CMS might in any case prefer to perform luminosity independent measurements and relate any of the “new” measurements to some relatively well known and measurable B-hadron decays.

It is also worth pointing out that currently no other high  $Q^2$  reaction has been envisioned, which might be measurable to a systematic precision of better than 5-10% and a luminosity of up to  $1\text{fb}^{-1}$ . In addition, most of the interesting high  $Q^2$  electroweak final states will unfortunately even be limited for the first few LHC years to statistical accuracies to 5% or more.

The prospect for the other luminosity measurements start to become at earliest interesting only once a few  $100\text{ pb}^{-1}$  can be recorded. Consequently one can expect to obtain a statistical interesting accuracy from the reaction  $pp \rightarrow pp\mu\mu$  after 2010. Similar, it looks unlikely that low luminosity special optics run will be performed before 2011. Consequently one might hope that few % accurate total cross section numbers become available before the 2012 data taking period will start.

## 6 Summary and Outlook

A large variety of potentially interesting pp luminosity measurements, proposed during the past 10-15 years, are presented in this Section.

Realistically only the machine luminosity measurement and the counting of the  $Z$  production might reach interesting accuracies of 5% before 2011. For all practical purposes it looks that both methods should be prepared in great detail before the data taking at 14 TeV collision energies will start in 2010.

We believe that a working group, consisting of interested members of the three pp collider experiments and interested theorists, should be formed to prepare the necessary Monte Carlo tools to make the best possible use of the soon expected  $W$  and  $Z$  data, not only for the pp luminosity normalization but even more for the detailed investigations of the parton parton luminosity determination and their use to predict other event rates for diboson production processes and high mass Drell-Yan events.

## References

- [1] Up to Date Performance of the ATLAS, CMS and LHCb Detectors and further detailed references can be found on the corresponding homepages <http://atlas.web.cern.ch/Atlas/index.html>, <http://cmsinfo.cern.ch/Welcome.html/> and <http://lhcb.web.cern.ch/lhcb/>.

- [2] W. Herr and B. Muratori, *Concept of luminosity*. Proceedings CAS2003, <http://doc.cern.ch/yellowrep/2006/p361.pdf>CERN-2006-002 p. 361.
- [3] H. Burkhardt and P. Grafstrom, *Absolute luminosity from machine parameters*. 2007. CERN-LHC-PROJECT-Report-1019 and <http://cdsweb.cern.ch/record/1056691>.
- [4] V. M. Budnev, I. F. Ginzburg, G. V. Meledin, and V. G. Serbo, Nucl. Phys. **B63**, 519 (1973).
- [5] V. A. Khoze, A. D. Martin, R. Orava, and M. G. Ryskin, Eur. Phys. J. **C19**, 313 (2001). [hep-ph/0010163](http://arxiv.org/abs/hep-ph/0010163).
- [6] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint [hep-ph/0511119](http://arxiv.org/abs/hep-ph/0511119), 2005.
- [7] CDF II Collaboration, D. E. Acosta *et al.*, Phys. Rev. Lett. **94**, 091803 (2005). [hep-ex/0406078](http://arxiv.org/abs/hep-ex/0406078).
- [8] A. M. Bellavance, *W/Z production cross sections and asymmetries at  $E(CM) = 2\text{-TeV}$* . Preprint [hep-ex/0506025](http://arxiv.org/abs/hep-ex/0506025), 2005.
- [9] F. P. M. Dittmar and D. Zurcher, Phys. Rev. **D56**, 7284 (1997). [hep-ex/9705004](http://arxiv.org/abs/hep-ex/9705004).
- [10] J. Pinfold, *Plans for the very forward region of ATLAS: The lucid luminosity monitor*. Prepared for 9th ICATPP Conference on Astroparticle, Particle, Space Physics, Detectors and Medical Physics Applications, Villa Erba, Como, Italy, 17-21 Oct 2005.
- [11] TOTEM Collaboration, G. Anelli *et al.*, JINST **3**, S08007 (2008).
- [12] LHCb Technical Proposal, CERN-LHCC-98-004.
- [13] LHCb Technical Design Report 9 *LHCb Reoptimized Detector*, CERN-LHCC-2003-030.
- [14] ATLAS Collaboration, N. Besson, M. Boonekamp, E. Klinkby, T. Petersen, and S. Mehlhase, Eur. Phys. J. **C57**, 627 (2008). [0805.2093](http://arxiv.org/abs/0805.2093).
- [15] See CMS Collaboration CMS PAS EWK-08-005 for Electrons and CMS PAS 2007/002 for Muons.
- [16] See LHCb Technical Design Report 10, *LHCb Trigger System*, CERN-LHCC-2003-031.
- [17] H. Schmickler, *How to measure beam intensity ?* <http://indico.cern.ch/getFile.py/access?contribId=s1t18&resId=1&materialId=1&confId=a053945>  
Talk given to Atlas 27 June 2005.
- [18] S. V. der Meer, *Calibration of the effective beam height in the isr*. <http://doc.cern.ch/archive/electronic/kek-scan//196800064.pdf>CERN-ISR-PO-.
- [19] K. Potter, *Luminosity measurements and calculations*. CAS 1992, CERN yellow report, <http://doc.cern.ch/yellowrep/1994/94-01/p117.pdf>. 117 ff.

- J. ANDERSON, M. BOONEKAMP, H. BURKHARDT, M. DITTMAR, V. HALYO, T. PETERSEN
- [20] G. C. *et al.*, Nucl. Phys. **B254**, 697 (1985).
- [21] M. Ferro-Luzzi, Nucl. Instrum. Meth. **A553**, 388 (2005).
- [22] The CMS Approach using Specialized High Rate Detectors is Described in the Talk of V.Halyo during the HERA-LHC Workshop “standard candles”.
- [23] The Presentations of the “standard candle” session during the HERA LHC Workshop in May 2008 can be found at:  
[http://indico.cern.ch/conferenceOtherViews.py?view=cdsagenda\\_olist&confId=27458#18..](http://indico.cern.ch/conferenceOtherViews.py?view=cdsagenda_olist&confId=27458#18..)
- [24] L3 Collaboration, P. Achard *et al.*, Phys. Lett. **B585**, 53 (2004). [hep-ex/0402037](http://arxiv.org/abs/hep-ex/0402037).
- [25] A. G. Shamov and V. I. Telnov, Nucl. Instrum. Meth. **A494**, 51 (2002).  
[hep-ex/0207095](http://arxiv.org/abs/hep-ex/0207095).
- [26] More Details about the LHCb Approach using  $pp \rightarrow pp^{+-}$ , are given in the talk of J. Anderson at the HERA-LHC Workshop “standard candles”.
- [27] K. M. C Anastasiou, L. J. Dixon and F. Petriello, Phys. Rev. **D69**, 094008 (2004).  
[hep-ph/0312266](http://arxiv.org/abs/hep-ph/0312266).
- [28] S. A. Y. N. E. Adam, V. Halyo, JHEP **05** (2008). 0802.3251.
- [29] S. . Y. W. Z. N. E. Ñadam, V.Ñhalyo, JHEP **09**, 13 (2008). 0808.0758.
- [30] R. S. Thorne, A. D. Martin, W. J. Stirling, and G. Watt, *Parton Distributions for the LHC*. Preprint 0706.0456, 2007.
- [31] LHCb Technical Design Report 10, *LHCb Trigger System*, CERN-LHCC-2003-031.
- [32] More details about the LHCb approach using  $\sigma_Z \cdot Br(Z \rightarrow \mu^+ \mu^-)$ , are given in the talk by J. Anderson at the HERA-LHC Workshop “standard candles”.
- [33] R. McNulty, *Potential PDF sensitivity at LHCb*. Preprint 0810.2550, 2008.
- [34] CMS collaboration, *Exclusive  $\gamma\gamma \rightarrow l^+ l^-$  and  $\gamma p \rightarrow \Upsilon p \rightarrow l^+ l^- p$  production*. CMS-PAS-DIF-07-001.

## Outlook: the PDF4LHC initiative

A. De Roeck

This document demonstrates the vast amount of progress that has taken place in the last years on pinning down the PDFs of the proton, as well as the dramatic increase in awareness of the impact of PDFs on the physics program of LHC experiments. The HERALHC workshop has acted as a regular forum for working meetings between the experiments, PDF phenomenologists and theorists. In the course of this workshop, it was realized that the momentum on the PDF studies should be kept and perhaps even focused more on the LHC, in order to continue the discussions, investigations and further work towards improving our knowledge on the PDFs.

Clearly, LHC will need the best PDFs, especially for precision measurements, setting of limits in searches, and even for discoveries. Ideally the ATLAS and CMS (and LHCb and ALICE) analyses should follow a common procedure for using PDFs and their uncertainties in their key analyses. Such a common procedure, across the experiments, is being used in other contexts, such as significance estimates in searches. Also, changing frequently the PDFs in the software of the experiments, e.g. for cross-checks or the determination of error bands, is often non-trivial (e.g. due to the inter-connection with parameter choices for underlying event modeling, showering parameters and so on) and sometimes impractical if CPU intensive detector simulations are involved. LHC studies therefore will need both good central values for the PDFs to start with, and a good estimate of the associated uncertainties.

This has triggered the so called PDF4LHC initiative. PDF4LHC offers a discussion forum for PDF studies and information exchange between all stake-holders in the field: the PDF global fitter groups, such as CTEQ and MSTW; the current experiments, such as the HERA and Tevatron ones; QCD theorists and the LHC experimental community. The PDF4LHC initiative started in 2008. More details and links to the meetings so far can be found on the PDF4LHC web site [1].

The mission statement of PDF4LHC is:

- Getting the best PDFs, including the PDF uncertainties, based on the present data.
- Devise strategies to use future LHC data to improve the PDFs.

All this needs a close collaboration between theorists and those that are preparing to make the measurements. In order to reach the first goal, the PDF4LHC forum aims to stimulate discussions and trigger further comparison exercises across the PDF community, in order to select one or a limited number of possible strategies that can be adapted to determine and use PDFs. For the second goal, PDF4LHC should also be a forum for discussions on how to include measurements from the LHC to constrain PDFs: what should be measured at LHC, and correspondingly calculated in theory. Such measurements include  $W$  and  $Z$  production and asymmetries, di-jet production, hard prompt photons, Drell-Yan production, bottom and top quark production,  $Z$ -shape fits and  $Z$ +jets measurements. One expects that some of these channels can already be studied with first data, hence we need to prepare for that well in advance.

The following issues are part of the program for in depth discussions via topical workshops, some of which took place already in 2008 [1].

- Data to be included in the PDFs. Would we get better results with a selection of data to be used? New data will become available such as  $F_L(x, Q^2)$ , and combined data from H1/ZEUS. Can we extract more from the data?

- Determination of PDF uncertainties, including the statistical treatment of the data.
- Theoretical uncertainties and regions/processes where they matter: higher-order corrections; heavy flavour treatment; low- $x$  (and high- $x$ ) resummation; other PDFs like unintegrated PDFs (and GPDs).
- PDFs for usage Monte Carlo generators.

One can expect that the LHC experiments most likely will be using for most of their studies the PDF sets and errors that are delivered by either one of the CTEQ or MSTW family. Hence it is important that the lessons learned from exercises on studies of the systematics on PDFs will be adapted by these main global PDF providers. PDF4LHC aims to advice the experiments in the use for PDFs for the LHC, based on the discussions, results and future consensus at the forum. The experience and results from HERAPDFs, and PDFs from other groups, like the Neural Net or Alekhin ones are extremely valuable in this discussion and will serve as crucial input in studies to demonstrate how well we actually know the parton distributions. Several important benchmark exercises have been already performed and are reported in section 3 of this report.

A special case are the PDFs for Monte Carlo generators. For experiments it is important that generated events be kinematically distributed close to the distribution of the real data, such that the simulated and reconstructed Monte Carlo events can be used in a straightforward way to calculate efficiencies for e.g. experimental cuts in an analysis. In case the initially generated distribution does not resemble the data close enough, the Monte Carlo samples need to be reweighted, with all its possible drawbacks. Since calculations based on LO Matrix Elements and LO PDFs are known not to describe the data well, and NLO Matrix Element based generators to date have so far only a restricted number of processes implemented, studies are ongoing on so called “improved LO” PDFs, which try to cure some of the LO PDF drawbacks. Examples are given in [2]. This is yet another part of the discussions in the PDF4LHC forum

In short, it is crucial that the work started here continues, with discussions and studies on PDFs and their uncertainties, the impact of the upcoming data on future PDF determinations and more, all with special focus on the needs for the LHC. The PDF4LHC initiative will offer a framework to do all this.

## References

- [1] The web page of the PDF4LHC forum can be found at <http://www.hep.ucl.ac.uk/pdf4lhc/>.
- [2] A. Sherstnev and R. S. Thorne, *Different PDF approximations useful for LO Monte Carlo generators*. Preprint 0807.2132, 2008.

## Chapter 2

### Working Group

# Multi-Jet Final States and Energy Flows

#### Convenors:

*C. Gwenlan (UCL, ZEUS)*

*L. Lönnblad (Lund),*

*E. Rodrigues (LHCb),*

*G. Zanderighi (CERN),*

*Contactpersons: S. Banerjee (CMS), D. Traynor (H1)*



## Working group: Multi-Jet final states and energy flows

Claire Gwenlan<sup>1</sup>, Leif Lönnblad<sup>2</sup>, Eduardo Rodrigues<sup>3</sup>, Giulia Zanderighi<sup>1</sup> (Eds.),  
Alessandro Bacchetta<sup>22</sup>, Andrea Banfi<sup>4</sup>, Sergey Baranov<sup>5</sup>, Jochen Bartels<sup>6</sup>,  
Armen Bunyatyan<sup>7,8</sup>, Victor Coco<sup>9</sup>, Gennaro Corcella<sup>10,11</sup>, Mrinal Dasgupta<sup>12</sup>, Michal Deák<sup>13</sup>,  
Pierre-Antoine Delsart<sup>9</sup>, Igor M. Dremin<sup>5</sup>, Francesco Hautmann<sup>1</sup>, S. Joseph<sup>14</sup>, Hannes Jung<sup>13</sup>,  
Albert Knutsson<sup>13</sup>, Krzysztof Kutak<sup>13</sup>, Artem Lipatov<sup>15</sup>, Gionata Luisoni<sup>16</sup>, Swapan Majhi<sup>14</sup>,  
Lluís Martí<sup>13</sup>, Katharina Müller<sup>17</sup>, Tim Namsoo<sup>13</sup>, Sakar Osman<sup>2</sup>, Hanno Perrey<sup>6</sup>,  
Germán Rodrigo<sup>18</sup>, Juan Rojo<sup>19</sup>, Zuzana Růriková<sup>13</sup>, Agustín Sabio Vera<sup>20</sup>, Christian Sander<sup>6</sup>,  
Thomas Schörner-Sadenius<sup>6</sup>, Florian Schwennsen<sup>21</sup>, Gábor Somogyi<sup>22</sup>, Grégory Soyez<sup>23</sup>,  
Mark Strikman<sup>24</sup>, Michele Treccani<sup>25,26</sup>, Daniele Treleani<sup>27</sup>, Zoltán Trócsányi<sup>28</sup>, B.F.L. Ward<sup>14</sup>,  
S.A. Yost<sup>29</sup>, Nikolai Zotov<sup>15</sup>

<sup>1</sup> Oxford University, UK; <sup>2</sup>Lund University, Sweden; <sup>3</sup> University of Glasgow, UK; <sup>22</sup>Jefferson Lab, USA; <sup>4</sup>Università di Milano–Bicocca, Italy; <sup>5</sup>Lebedev Physical Institute, Moscow, Russia; <sup>6</sup>University of Hamburg, Germany; <sup>7</sup>MPI-K, Heidelberg Germany; <sup>8</sup>YerPhI, Yerevan, Armenia; <sup>9</sup>LAPP, Annecy, France; <sup>10</sup>Museo Storico della Fisica e Centro Studi e Ricerche E. Fermi, Italy; <sup>11</sup>Scuola Normale Superiore Italy; <sup>12</sup>University of Manchester, UK; <sup>13</sup>DESY, Hamburg, Germany; <sup>14</sup> Baylor University, Waco, USA; <sup>15</sup>Skobeltsyn Institute of Nuclear Physics, Lomonosow Moscow State University, Russia.; <sup>16</sup>Universität Zürich, Switzerland; <sup>17</sup>University of Zurich, Switzerland; <sup>18</sup>IFIC, CSIC-Universitat de València, Spain; <sup>19</sup>LPTHE, Paris, France; <sup>20</sup>CERN, Geneva, Switzerland & Instituto de Física Teórica UAM/CSIC, Universidad, Autónoma de Madrid, E-28049 Madrid, Spain <sup>21</sup>LPT Université d’Orsay, CNRS, France; <sup>22</sup>University of Zürich, Switzerland; <sup>23</sup>Brookhaven National Laboratory, USA; <sup>24</sup>Pennsylvania State University, USA; <sup>25</sup>Università di Ferrara, Italy; <sup>26</sup>Universidad de Granada, Spain; <sup>27</sup>Università di Trieste, Italy; <sup>28</sup>University of Debrecen, Hungary; <sup>29</sup>The Citadel, Charleston, USA.

### Abstract

We present a summary of the activities of the *Multi-Jet final states and energy flows* Working Group of the *HERA and the LHC workshop*, 2007-2008. Among the more specific topics considered were the status of and recent progress in higher order calculations, both in fixed perturbative expansions and in resummed approaches, recent progress in the description of jets, including the description of forward jets, new calculations performed using  $k_T$ -factorization and new determinations of unintegrated parton densities.

## 1 Introduction

*Authors: Claire Gwenlan, Leif Lönnblad, Eduardo Rodrigues, Giulia Zanderighi*

The activities of Working Group 2, *Multi-Jet Final States and Energy Flows*, have covered a broad range of topics, encompassing both theoretical and experimental advances in understanding the hadronic final state at high energies. Much of this work will be of significant benefit in preparing to fully exploit the LHC physics potential. We focus here on progress in the field since the last proceedings of this workshop [1, 2].

From a theoretical point of view, a good understanding of the Standard Model (SM) is of the utmost importance in order to be able to unravel and disentangle possible New Physics effects. In addition, the study of the Standard Model is important in its own right, especially in the QCD sector where the strong coupling in many cases prevents us from making reliable predictions. Recently, considerable progress has been made in the area of higher order calculations in perturbative QCD. Some developments are discussed in the contributions of Sec. [3]. Typically, these fixed-order calculations are sufficient to describe inclusive observables, such as cross sections or transverse momentum ( $p_t$ ) spectra at sufficiently high scales. However, more exclusive observables, such as event-shape distributions, require that one rearranges the perturbative expansion and that one resums leading and next-to-leading logarithmic terms to all orders in perturbation theory. This technology is today already well-developed both in terms of analytical calculations and in terms of numerical implementations in parton shower Monte Carlos. We report on further recent progress in the understanding and development of such resummed calculations in Sec. [4].

The development and use of jet algorithms plays a key role in the study of hadronic final states. Indeed jets are an essential tool for a variety of studies, such as top reconstruction, mass measurements and searches for Higgs and new physics. Furthermore, they are instrumental for QCD studies, e. g. for inclusive-jet measurements, which in turn constitute an important input for parton density determinations. By clustering particles into jets, jet algorithms reduce complicated multi-particle events in simple final states with few jets. This procedure and the way particles are recombined together (e.g. the  $E$ - or  $P$ -scheme) is fundamentally non-unique. This freedom can be exploited to extract information from jets. The rapid, recent development of fast, infrared- and collinear-safe cone and clustering algorithms, is discussed in Sec. [5]. Also considered are the issues of jet-finding, reconstruction and calibration currently being developed by the LHC experimental collaborations. Recent work on defining jet-quality measures, designed to quantify the performance of jet algorithms, is also presented.

In Sec. [6] we focus our attention to the  $k_T$ -factorization approach, which may be the key to fully understand the hadronic final states at the LHC. Although the standard collinear factorization should hold for the description of jets at very high scales, we expect it to break down at somewhat smaller scales and low  $x$ , and the use of  $k_T$ -factorization and unintegrated parton densities will become essential. This is an area where we have learned a lot from HERA results, and where we may learn more still from data yet to be analyzed.

A major difficulty in describing final states at high energies is the treatment of multi-parton interactions. There is no doubt that, due to the high density of small- $x$  partons, the events at the LHC will contain several semi-hard parton-parton scatterings. Indeed such events have already been studied at the Tevatron, and models including this feature are needed in order to describe e.g. the underlying events in photo-production at HERA. Although models for multi-

parton interactions exist, there are many uncertainties, and the differences in the predictions for the LHC are large. Most of the work on multi-parton interactions in the workshop was presented in joint sessions with the Monte Carlo tools working group, and the corresponding contribution to these proceedings are presented in the section of this working group [7].

In Sec. [8], reviews some recent experimental results from HERA which are of interest for future LHC studies, concentrating particularly on isolated photon and jet production, including the effects of multi-parton interactions and the underlying event.

Finally we take a look at processes at even higher parton densities, such as those occurring in heavy ion collisions. Here it is important to consider not only the productions of jets, and possible effects of gluon saturation, but also the propagation of the hard partons through a dense medium. A couple of issues related to such interactions at high densities are discussed in Sec. [9].

### Acknowledgments

G. Rodrigo acknowledges support by MICINN under Grant No. FPA2007-60323 and No. INFN08-02, and by Generalitat Valenciana under Grant No. PROMETEO/2008/069

### References

- [1] S. Alekhin *et al.* (2005), arXiv:hep-ph/0601012.
- [2] S. Alekhin *et al.* (2005), arXiv:hep-ph/0601013.
- [3] G. Zanderighi *et al.*, *Higher-order calculations*. These proceedings.
- [4] A. Banfi *et al.*, *Event shapes and resummation*. These proceedings.
- [5] V. Coco *et al.*, *Jets and jet algorithms*. These proceedings.
- [6] S. Baranov *et al.*,  *$k_{\perp}$ -factorization and forward jets*. These proceedings.
- [7] S. Gieseke *et al.*, *Working group summary — monte carlo tools*. These proceedings.
- [8] K. Müeller *et al.*, *Hera results*. These proceedings.
- [9] M. Strikman and I. Dremin, *Interactions at high gluon densities*. These proceedings.

## Higher-order calculations

*Giulia Zanderighi, Germán Rodrigo, Michele Treccani, Gábor Somogyi*

The start-up of the LHC will usher in a new era of discovery in high-energy physics, with the machine operating at the highest centre-of-mass energy ever attained in the laboratory. In order to fully exploit its physics potential in Higgs and beyond the Standard Model (BSM) searches, a good understanding of the Standard Model is necessary. This requires a precise theoretical understanding of QCD.

The simplest description in exact perturbative calculations is at leading order (LO) using collinear factorization. Here, partons (or particles) should be well-separated and hard so as to avoid large soft-collinear corrections. Today, these LO calculations are fully automated. However, the drawback is that they have very large scale dependencies, enhanced sensitivities to kinematical cuts and a poor modelling of the jet structure (each parton corresponding to a jet). Therefore it is currently well appreciated that accurate predictions of QCD jet cross sections require the computation of radiative corrections at least to next-to-leading order (NLO) accuracy, first for SM processes, and BSM processes at a later stage. This is simply because the QCD coupling is not small and the phase space for emitting additional partons at the LHC is large, so that NLO corrections can be numerically significant. Benefits of NLO include a reduced dependence on unphysical scales, a better modelling of jets, and a more reliable control of the normalization and shape of cross sections.

Three ingredients are needed to compute a  $2 \rightarrow N$  process at NLO: the real radiation of one parton from the  $2 + N$  parton system (tree-level  $2 + N + 1$  processes), one-loop virtual corrections to the  $2 \rightarrow N$  process and a method to cancel the divergences of real and virtual corrections before numerical integration. The calculation of tree-level amplitudes has been automated and also the cancellation of divergences is, today, well understood [1–3]. Therefore up until very recently, the bottleneck at NLO has been the calculation of virtual, loop amplitudes.

In some cases however, NLO accuracy is not yet satisfactory and one would like to be able to calculate perturbative corrections beyond NLO. The physical situations when this happens have been discussed extensively in the literature [4]. Usually NLO is insufficient when the NLO correction is comparable to, or larger than, the LO result. This may happen when a process involves very different scales, so that large logarithms of the ratio of the two scales arise, which need to be resummed. This may also happen when new channels open up (at NLO those channels are effectively LO). This is the case, for instance, for  $b$ -jet production, where gluon splitting and flavour excitation processes enter at NLO and are enhanced by large logarithms. Also, gluon dominated processes are often characterized by large corrections, both because gluons radiate on average more than quarks and because of the steeply falling parton distribution functions (PDFs) at small  $x$ . NLO might also be insufficient if very high precision is useful. This is occasionally the case, for instance, in Drell-Yan processes, top pair production, and 3-jet production in  $e^+e^-$ . Finally, since NLO provides a first reliable estimate of cross sections, only NNLO can in principle provide a reliable error estimate of those cross sections. The bottleneck at NNLO is not the calculation of virtual matrix elements, as is the case at NLO, but rather the cancellation of divergences before numerical evaluation. In the following we will report on some recent progress in higher-order perturbative QCD.

## 1 One-loop amplitudes: the gluon case

*Author: Giulia Zanderighi*

Current and upcoming collider experiments require a good understanding of Standard Model (SM) processes in order to carry out any successful search for a Higgs or beyond SM signals (BSM). Therefore, these searches will benefit from next-to-leading order predictions, for SM processes first, and BSM processes at a later stage. Traditional Feynman diagram techniques, supplemented by robust numerical methods (Passarino-Veltman decomposition, Davydychev reduction, integration by part, tensor reduction) are well developed and made it possible to develop powerful computation tools [5–8] including procedures to handle potential numerical instabilities [7,9]. These techniques have been applied recently in a variety of  $2 \rightarrow 3$  scattering processes and pushed to their limit in few  $2 \rightarrow 4$  cases (see [10] for a recent review). The bottleneck of these approaches is the rapid increase both in the number of Feynman diagrams and in the number of terms generated during the tensors reduction. One promising alternative method is based on generalized unitarity [11]. Recent advances [12] allowed the development of analytic methods for the calculation of the full amplitude, including the rational part, using recursion relations [13,14]. A recent computational scheme is based on unitarity in integer higher dimension [15,16]. This allows one to reduce the calculation of *full* one-loop amplitudes to the calculation of residues and of tree-level amplitudes involving complex momenta.

Using unitarity in higher integer dimension together with Berends-Giele recursion relations, we show that it is possible to develop an algorithm of mild, polynomial complexity for the evaluation of one-loop amplitudes. As a first application, we considered here pure gluonic amplitudes. We analyze the numerical stability of the results and the time dependence of the algorithm for virtual amplitudes with up to twenty external gluons.

### 1.1 The method

We [17] implemented the methods developed in Refs. [15,16] with some minor modifications into the **Rocket** program. These methods build upon the formalism of Ref. [18] by removing the requirement of the four dimensional spinor language, thereby allowing for the extension of the method to  $D$ -dimensional cuts. To calculate the full one-loop  $N$ -gluon amplitude, it is sufficient to be able to calculate the leading colour ordered one-loop amplitude, since from these colour ordered amplitudes the full one-loop amplitude can be constructed [11,19]. In the following we will therefore focus on the leading colour ordered amplitudes  $A_N^{[1]}(1, 2, \dots, N)$ . We will use the (over-complete) master integral basis decomposition derived in Ref. [16]

$$\begin{aligned}
A_N^{[1]} &= - \sum_{[i_1|i_5]} \frac{(D-4)}{2} c_{i_1 i_2 i_3 i_4 i_5}^{(2,0)} I_{i_1 i_2 i_3 i_4 i_5}^{(D+2)} \quad (1) \\
&+ \sum_{[i_1|i_4]} \left( d_{i_1 i_2 i_3 i_4}^{(0,0)} I_{i_1 i_2 i_3 i_4}^{(D)} - \frac{(D-4)}{2} d_{i_1 i_2 i_3 i_4}^{(2,0)} I_{i_1 i_2 i_3 i_4}^{(D+2)} + \frac{(D-4)(D-2)}{4} d_{i_1 i_2 i_3 i_4}^{(4,0)} I_{i_1 i_2 i_3 i_4}^{(D+4)} \right) \\
&+ \sum_{[i_1|i_3]} \left( c_{i_1 i_2 i_3}^{(0,0)} I_{i_1 i_2 i_3}^{(D)} - \frac{(D-4)}{2} c_{i_1 i_2 i_3}^{(2,0)} I_{i_1 i_2 i_3}^{(D+2)} \right) \\
&+ \sum_{[i_1|i_2]} \left( b_{i_1 i_2}^{(0,0)} I_{i_1 i_2}^{(D)} - \frac{(D-4)}{2} b_{i_1 i_2}^{(2,0)} I_{i_1 i_2}^{(D+2)} \right),
\end{aligned}$$

where we introduced the short-hand notation  $[i_1|i_n] = 1 \leq i_1 < i_2 < \dots < i_n \leq N$  and

$$I_{i_1, \dots, i_N}^D = \int \frac{d^D l}{i\pi^{D/2}} \frac{1}{d_{i_1} d_{i_2} \dots d_{i_N}}, \quad d_i = d_i(l) = (l + q_i)^2 = (l + p_1 + \dots + p_i)^2. \quad (2)$$

Because some coefficients are multiplied with a dimensional factor  $(D-4)$  they cannot be determined using four dimensional cuts, therefore we extend the dimensionality of the cut line to integer, higher dimensions, resulting in a well-defined on-shell particle after performing the cut [16]. By applying quintuple, quadruple, triple and double  $D_s$ -dimensional cuts (where  $D_s \geq D$  denotes the dimensionality of the spin-space) we can determine the coefficients of the parametric form of the one-loop amplitude. This requires the calculation of the factorized unintegrated one-loop amplitude

$$\begin{aligned}
\text{Res}_{i_1 \dots i_M}(\mathcal{A}_N^{[1]}(l)) &= \left( d_{i_1} \times \dots \times d_{i_M} \times \mathcal{A}_N^{[1]}(l) \right)_{d_{i_1} = \dots = d_{i_M} = 0} \\
&= \sum_{\{\lambda_1, \dots, \lambda_M\}=1}^{D_s-2} \left( \prod_{k=1}^M \mathcal{A}_{i_{k+1}-i_k}^{[0]}(l_{i_k}^{(\lambda_k)}, p_{i_k+1}, \dots, p_{i_{k+1}}, -l_{i_{k+1}}^{(\lambda_{k+1})}) \right), \quad (3)
\end{aligned}$$

where  $M \leq 5$  and the  $D$ -dimensional loop momentum  $l$  has to be chosen such that  $d_{i_1}(l) = \dots = d_{i_M}(l) = 0$ . To calculate these tree amplitudes we use the standard Berends-Giele recursion relation [20] which is valid in arbitrary dimension and for complex momenta. The generic solution for the loop momentum in Eq. (3) is given by

$$l_{i_1 \dots i_M}^\mu = V_{i_1 \dots i_M}^\mu + \sqrt{\frac{-V_{i_1 \dots i_M}^2}{\alpha_M^2 + \dots + \alpha_D^2}} \left( \sum_{i=M}^D \alpha_i n_i^\mu \right), \quad (4)$$

for arbitrary values of the variables  $\alpha_i$ . The vector  $V_{i_1 \dots i_m}^\mu$  is defined in the space spanned by the denominator offset momenta  $\{q_{i_1}, \dots, q_{i_m}\}$ , while the orthonormal basis vectors  $\{n_M^\mu, \dots, n_D^\mu\}$  span the space orthogonal to the space spanned by these momenta [15, 16]. Given the solution to the on-shell conditions  $l_{i_1 \dots i_M}^\mu$  in Eq. (4), the loop momenta flowing into the tree amplitudes  $l_{i_k}$  and  $l_{i_{k+1}}$  in Eq. (3) are fixed by momentum conservation (see Ref. [15]). Once all coefficients in Eq. (1) have been determined we can continue the dimensionality to the non-integer limit:

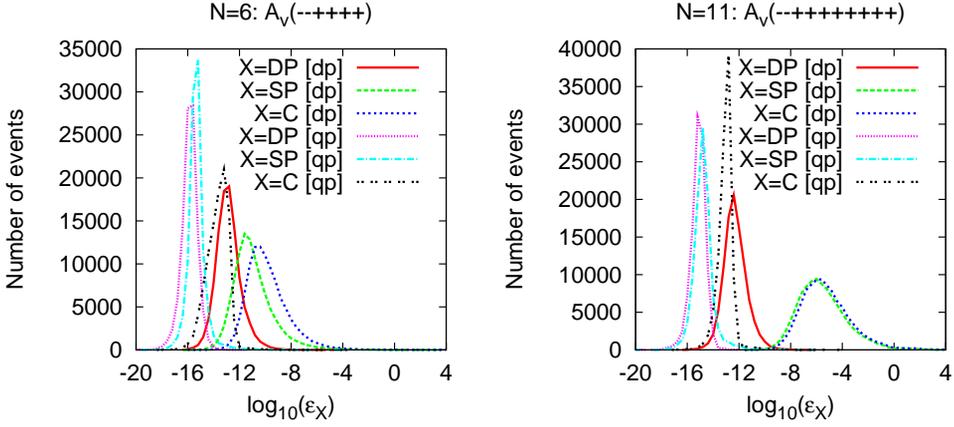


Fig. 1: Accuracy on the double pole, single pole and constant part of the maximally helicity violating (MHV) amplitude with adjacent negative helicities for 6 up to 11 external gluons. Double ([dp]) and quadrupole ([qp]) precision results for 100,000 phase space points are shown. Refer to the text for more details.

$D \rightarrow 4 - 2\epsilon$ . Neglecting terms of order  $\epsilon$  we find for the colour ordered one-loop amplitude

$$\begin{aligned}
 A_N^{[1]} &= \sum_{[i_1|i_4]} d_{i_1 i_2 i_3 i_4}^{(0,0)} I_{i_1 i_2 i_3 i_4}^{(4-2\epsilon)} + \sum_{[i_1|i_3]} c_{i_1 i_2 i_3}^{(0,0)} I_{i_1 i_2 i_3}^{(4-2\epsilon)} + \sum_{[i_1|i_2]} b_{i_1 i_2}^{(0,0)} I_{i_1 i_2}^{(4-2\epsilon)} \\
 &- \sum_{[i_1|i_4]} \frac{d_{i_1 i_2 i_3 i_4}^{(4,0)}}{6} + \sum_{[i_1|i_3]} \frac{c_{i_1 i_2 i_3}^{(2,0)}}{2} - \sum_{[i_1|i_2]} \frac{(q_{i_1} - q_{i_2})^2}{6} b_{i_1 i_2}^{(2,0)} + \mathcal{O}(\epsilon). \quad (5)
 \end{aligned}$$

The terms in the first line give rise to the so-called cut-constructable part of the amplitude [21]. The terms in the second line can be identified with the rational part. In the approach used here the division between these two contributions is irrelevant. For the numerical evaluation of the bubble, triangle and box master integrals we use the package developed in Ref. [22].

## 1.2 Numerical results: accuracy and time dependence of the algorithm

To study the numerical accuracy of the on-shell method implemented in **Rocket** we define

$$\epsilon_C = \log_{10} \frac{|A_N^{v,\text{unit}} - A_N^{v,\text{analy}}|}{|A_N^{v,\text{analy}}|}, \quad (6)$$

where “unit” denotes the result obtained with the on-shell method and “analy” the analytical result for the constant parts of the one-loop helicity amplitudes (or in the case of  $N = 6$  the numerical results of [23]). Similarly, we denote by  $\epsilon_{\text{DP}}$  and  $\epsilon_{\text{SP}}$  the accuracy on the double and single poles, respectively.

In Fig. (1) we show the accuracy for the two adjacent minus helicity gluon MHV one-loop amplitudes,  $A_N^{[1]}(--+\dots+)$ , for  $N = 6$  and  $N = 11$ , which are known analytically [11,21,24]. The 100,000 phase space points used for each multiplicity are generated uniformly in phase space

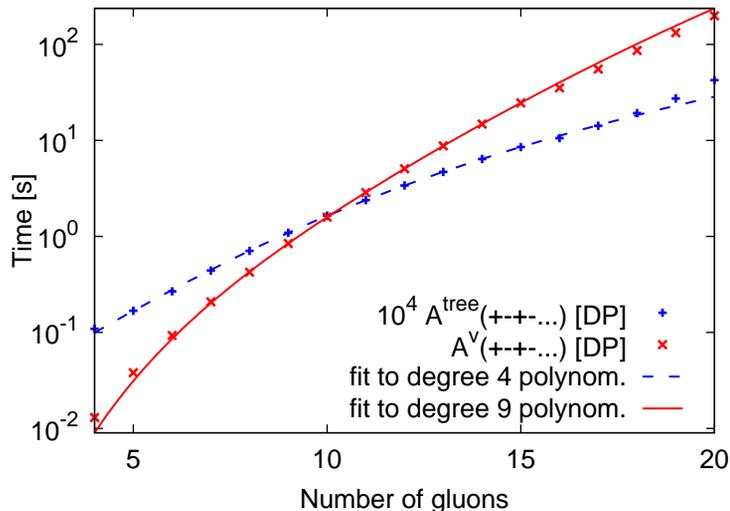


Fig. 2: Time in seconds needed to compute tree (blue, dashed) and one-loop (red, solid) ordered amplitudes with gluons of alternating helicity signs,  $A_N^{[1]}(+ - + - + \dots)$ , as a function of the number of external gluons ranging between 4 to 20 using a single 2.33 GHz Xeon processor.

using the Rambo algorithm [25] imposing minimal cuts. We plot the accuracy for the double pole ( $X = \text{DP}[\text{dp}]$ , solid, red), the single pole ( $X = \text{SP}[\text{dp}]$ , green, dot-dashed) and the constant part ( $X = \text{C}[\text{dp}]$ , blue, dotted). We see that an excellent accuracy can be reached for all contributions. The tail of the distribution reaching to large values of  $\epsilon$  contains only a very few points. This lack of agreement is due to numerical instabilities due to vanishing Gram determinants or other small intermediate denominators. Several techniques have been developed to deal with such exceptional points, such as developing systematic expansions [6, 7, 9] or interpolating across the singular regions [26]. We adopt here a more brute force approach and recur to quadrupole precision. In Fig. (1), we see three more curves marked [qp]: they correspond to the numerical accuracy on the same phase space points when the one-loop amplitude is computed in quadrupole precision. Out of 100,000 phase space points sampled, not a single one has an accuracy worse than  $10^{-4}$  and, at quadrupole precision we see no appreciable worsening of the accuracy with increasing  $N$ . Therefore up to  $N = 11$  (and probably even for more gluons) quadrupole precision is sufficient to guarantee an accuracy needed for any next-to-leading order QCD correction. If higher precision is desired one can choose to evaluate the few phase space points which have insufficient precision using some arbitrary precision package, at the cost of higher computation time. We note that while the plots here presented are for the MHV amplitudes, we performed a similar study for the finite amplitudes ( $A_N^{[1]}(+ \dots +)$ ,  $A_N^{[1]}(- + \dots +)$ ) and obtain very similar results. This indicates that the accuracy is essentially independent of the helicities of the external gluons.

A very important property of this method is that the time needed to compute one-loop amplitudes does not grow factorially with the number of external legs; indeed it is straightforward to estimate the scaling of time with the number of gluons  $N$ . The calculation of tree-level am-

plitudes computed via Berends-Giele recursion relations with caching of previously computed amplitudes requires a time which grows as  $\tau_{\text{tree},N} \propto N^4$  [27]. The total number of tree amplitudes that one needs to evaluate to get a one-loop amplitudes is given by

$$n_{\text{tree}} = \{(D_{s_1} - 2)^2 + (D_{s_2} - 2)^2\} \times \left( 5 c_{5,\text{max}} \binom{N}{5} + 4 c_{4,\text{max}} \binom{N}{4} + 3 c_{3,\text{max}} \binom{N}{3} + 2 c_{2,\text{max}} \left[ \binom{N}{2} - N \right] \right), \quad (7)$$

where the first factor is due to the sum over polarization of the internal cut gluons in two integer dimensions  $D_{s_1}$  and  $D_{s_2}$ . The constants  $c_{m,\text{max}}$  denote the number of times one needs to perform a multiple cut in order to fully constrain the system of equations determining the master integral coefficients. Explicitly one has  $c_{5,\text{max}} = 1$ ,  $c_{4,\text{max}} = 5$ ,  $c_{3,\text{max}} = 10$ , and  $c_{2,\text{max}} = 10$ . The integer number in front counts the number of tree amplitudes per multiple cut, finally the binomial coefficients corresponds to the number of possible cuts (for two point functions we subtract the vanishing contributions of the external self energy graphs). It follows that the time needed to evaluate a one-loop ordered amplitude will for large  $N$  scale as

$$\tau_{\text{one-loop},N} \sim n_{\text{tree}} \cdot \tau_{\text{tree},N} \propto N^9. \quad (8)$$

In Fig. (2) we plot the time needed to compute tree (blue, dashed) and one-loop (red, solid) ordered amplitudes with alternating helicity signs for the gluons,  $A_N^{[1]}(+ - + - \dots)$ , as a function of the number of gluons in the range between four and twenty. Time estimates refer to using a 2.33 GHz Xeon processor. One can see that the times needed to compute tree and one-loop ordered amplitudes are consistent with a  $N^4$  and  $N^9$  growth respectively. When running in quadrupole precision rather than in double precision the evaluation time grows, but the scaling with  $N$  remains unchanged. Finally we remark that the time is independent on the helicities of the external gluons.

### 1.3 Discussion and outlook

The results presented here are based on  $D$ -dimensional unitarity implemented in the Fortran 90 code **Rocket**. The very mild, power-like increase in computational time and the numerical stability of the results demonstrate the power of this approach. The large number of gluons considered here demonstrates that the gluon case is fully solved as far as virtual amplitudes are concerned.

Recently this method has been applied also to other processes  $0 \rightarrow t\bar{t}ggg$  [28],  $0 \rightarrow q\bar{q}W + n$  gluons and  $0 \rightarrow q\bar{q}\bar{Q}QW + 1$  gluon [29]. These recent calculations demonstrate the generality of the approach and constitute first steps towards automated one-loop calculations.

## 2 Duality relation between one-loop integrals and single-cut phase-space integrals

*Author: Germán Rodrigo*

As discussed in Sec. , the physics program of the LHC requires the evaluation of multi-leg signal and background processes at next-to-leading order (NLO). In the recent years, important efforts have been devoted to the calculation of many  $2 \rightarrow 3$  processes and some  $2 \rightarrow 4$  processes (see, e.g., [10]).

We have recently proposed a method [30–32] to numerically compute multi-leg one-loop cross sections in perturbative field theories. The starting point of the method is a duality relation between one-loop integrals and phase-space integrals. The duality relation requires to properly regularize propagators by a complex Lorentz-covariant prescription, which is different from the customary  $+i0$  prescription of the Feynman propagators. This duality relation has analogies with the Feynman’s Tree Theorem (FTT) [33], but involves only single cuts of the one-loop Feynman diagrams.

The duality relation between one-loop integrals and single-cut phase-space integrals is obtained [32] by applying the Cauchy residue theorem to a generic one-loop integral  $L^{(N)}$ :

$$L^{(N)}(p_1, p_2, \dots, p_N) = \int_q \prod_{i=1}^N G(q_i), \quad \int_q \dots \equiv -i \int \frac{d^d q}{(2\pi)^d} \dots, \quad (9)$$

where  $q_i = q + \sum_{k=1}^i p_k$  are the momenta of the internal lines, with  $q$  the loop momentum, and  $p_i$  ( $\sum_{i=1}^N p_i = 0$ ) the external (outgoing and clockwise ordered) momenta, and  $G$  is the customary Feynman propagator, which for massless internal lines is given by

$$G(q) \equiv \frac{1}{q^2 + i0}. \quad (10)$$

In the complex plane of the loop energy  $q_0$  the Feynman propagator has two poles; the pole with positive (negative) energy is slightly displaced below (above) the real axis. Hence, by using the Cauchy residue theorem in the  $q_0$  complex plane, with the integration contour closed at  $\infty$  in the lower half-plane, we obtain

$$L^{(N)}(p_1, p_2, \dots, p_N) = -2\pi i \int_{\mathbf{q}} \sum \text{Res}_{\{\text{Im } q_0 < 0\}} \left[ \prod_{i=1}^N G(q_i) \right]. \quad (11)$$

The Feynman propagators produce  $N$  poles in the lower half-plane that contribute to the residues in Eq. (11). The calculation of these residues is elementary, but it involves several subtleties. We get

$$\text{Res}_{\{i^{\text{th}} \text{ pole}\}} \frac{1}{q_i^2 + i0} = \int dq_0 \delta_+(q_i^2). \quad (12)$$

This result shows that considering the residue of the Feynman propagator of the internal line with momentum  $q_i$  is equivalent to cutting that line by including the corresponding on-shell propagator  $\delta_+(q_i^2)$ . The other propagators  $G(q_j)$ , with  $j \neq i$ , which are not singular at the value of the pole of  $G(q_i)$ , contribute as follows [32]:

$$\prod_{j \neq i} \frac{1}{q_j^2 + i0} \Big|_{q_i^2 = -i0} = \prod_{j \neq i} \frac{1}{q_j^2 - i0 \eta(q_j - q_i)}, \quad (13)$$

where  $\eta$  is a future-like vector, i.e. a  $d$ -dimensional vector that can be either light-like ( $\eta^2 = 0$ ) or time-like ( $\eta^2 > 0$ ) with positive definite energy ( $\eta_0 \geq 0$ ). The calculation of the residue at the pole of the  $i^{\text{th}}$  internal line modifies the  $i0$  prescription of the propagators of the other

internal lines of the loop. This modified regularization is named ‘dual’  $i0$  prescription, and the corresponding propagators are named ‘dual’ propagators. The dual prescription arises from the fact that the original Feynman propagator  $1/(q_j^2 + i0)$  is evaluated at the *complex* value of the loop momentum  $q$ , which is determined by the location of the pole at  $q_i^2 + i0 = 0$ . The presence of  $\eta$  is a consequence of the fact that the residue at each of the poles is not a Lorentz-invariant quantity, because a given system of coordinates has to be specified to apply the residue theorem. Different choices of the future-like vector  $\eta$  are equivalent to different choices of the coordinate system. The Lorentz-invariance of the loop integral is, however, recovered after summing over all the residues.

Inserting the results of Eqs. (12)-(13) in Eq. (11) gives us the duality relation between one-loop integrals and single-cut phase-space integrals [32]:

$$L^{(N)} = -\tilde{L}^{(N)}, \quad (14)$$

where the explicit expression of the phase-space integral  $\tilde{L}^{(N)}$  is

$$\tilde{L}^{(N)}(p_1, p_2, \dots, p_N) = \int_q \sum_{i=1}^N \tilde{\delta}(q_i) \prod_{\substack{j=1 \\ j \neq i}}^N \frac{1}{q_j^2 - i0 \eta(q_j - q_i)}, \quad (15)$$

with  $\tilde{\delta}(q) \equiv 2\pi i \delta_+(q^2)$ . Contrary to the FTT, the duality relation involves single-cut contributions only. This result is achieved by replacing the Feynman propagators in  $L^{(N)}$  by dual propagators in  $\tilde{L}^{(N)}$ , which depend on the auxiliary vector  $\eta$ . However,  $\tilde{L}^{(N)}$  does not depend on  $\eta$ , provided it is fixed to be the same in all its contributing single-cut terms (dual integrals). The duality relation, therefore, directly expresses the one-loop integral as the phase-space integral of a tree-level quantity. In the case of the FTT, the relation between loop and tree-level quantities is more involved, since the multiple-cut contributions contain integrals of expressions that correspond to the product of  $m$  tree-level diagrams over the phase-space for different number of particles.

The FTT and the duality theorem can be directly related starting from a basic identity between dual and Feynman propagators [32]:

$$\tilde{\delta}(q) \frac{1}{2qk + k^2 - i0 \eta k} = \tilde{\delta}(q) \left[ G(q+k) + \theta(\eta k) \tilde{\delta}(q+k) \right]. \quad (16)$$

This identity applies to the dual propagators when they are inserted in a single-cut integral. The proof of equivalence of the FTT and the duality theorem is purely algebraic [32]. We explicitly illustrate it by considering the massless two-point function  $L^{(2)}(p_1, p_2)$ . Its dual representation is

$$\tilde{L}^{(2)}(p_1, p_2) = \int_q \tilde{\delta}(q) \left( \frac{1}{2qp_1 + p_1^2 - i0 \eta p_1} + (p_1 \leftrightarrow p_2) \right). \quad (17)$$

Inserting Eq. (16) in Eq. (17), we obtain

$$\tilde{L}^{(2)}(p_1, p_2) = L_{1\text{-cut}}^{(2)}(p_1, p_2) + [\theta(\eta p_1) + \theta(\eta p_2)] L_{2\text{-cut}}^{(2)}(p_1, p_2), \quad (18)$$

where the  $m$ -cut integrals  $L_{m\text{-cut}}^{(2)}$  are the contributions with precisely  $m$  delta functions:

$$L_{1\text{-cut}}^{(2)}(p_1, p_2) = \int_q \tilde{\delta}(q) (G(q + p_1) + G(q + p_2)) , \quad L_{2\text{-cut}}^{(2)}(p_1, p_2) = \int_q \tilde{\delta}(q) \tilde{\delta}(q + p_1) . \quad (19)$$

Owing to momentum conservation (namely,  $p_1 + p_2 = 0$ ),  $\theta(\eta p_1) + \theta(\eta p_2) = 1$ , and then the dual and the FTT representations of the two-point function are equivalent. The proof of equivalence in the case of higher  $N$ -point functions proceeds in a similar way [32], the key ingredient simply being the constraint of *momentum conservation*.

The extension of the duality relation to include propagators with real finite masses  $M_i$  is straightforward. The massless on-shell delta function  $\tilde{\delta}(q_i)$  is replaced by  $\tilde{\delta}(q_i; M_i) = 2\pi i \delta_+(q_i^2 - M_i^2)$  when a massive loop internal line is cut to obtain the dual representation. The  $i0$  prescription of the dual propagators is not affected by real masses. The corresponding dual propagator is

$$\frac{1}{q_j^2 - M_j^2 - i0 \eta(q_j - q_i)} . \quad (20)$$

Unstable particles, in contrast, introduce a finite imaginary contribution in their propagators. The form of the complex-mass propagators is scheme dependent, but their poles in the  $q_0$  complex plane are located at a finite imaginary distance from the real axis. Then, when complex-mass propagators are cut in the duality relation, the  $+i0$  prescription of the usual Feynman propagators can be removed.

The polarization tensor of a spin-one gauge boson has in general the form

$$d^{\mu\nu}(q) = -g^{\mu\nu} + (\zeta - 1) \ell^{\mu\nu}(q) G_G(q) . \quad (21)$$

The second term on the right-hand side is absent only in the 't Hooft–Feynman gauge ( $\zeta = 1$ ). The tensor  $\ell^{\mu\nu}(q)$ , which propagates longitudinal polarizations, has a polynomial dependence on the momentum  $q$  and, therefore, it does not interfere with the residue theorem. The factor  $G_G(q)$  ('gauge-mode' propagator), however, can introduce extra unphysical poles (i.e. in addition to the poles of the associated Feynman propagator) that will modify the duality relation. Apart from the 't Hooft–Feynman gauge, the duality relation in the form presented here, i.e. with the inclusion of the sole single-cut terms from the Feynman propagators, turns out to be valid [32] in spontaneously-broken gauge theories in the unitary gauge, and in unbroken gauge theories in physical gauges specified by a gauge vector  $n^\nu$ , *provided* the dual vector  $\eta^\mu$  is chosen such that  $n \cdot \eta = 0$ . This excludes gauges where  $n^\nu$  is time-like. In any other gauge, additional single-cut terms from the absorptive contribution of the unphysical gauge poles have to be introduced in the duality relation.

The duality relation can be applied to evaluate not only basic one-loop integrals  $L^{(N)}$  but also complete one-loop quantities  $\mathcal{A}^{(1\text{-loop})}$  (such as Green's functions and scattering amplitudes). The analogue of Eqs. (14) and (15) is the following duality relation [32]:

$$\mathcal{A}^{(1\text{-loop})} = - \tilde{\mathcal{A}}^{(1\text{-loop})} . \quad (22)$$

The expression  $\tilde{\mathcal{A}}^{(1\text{-loop})}$  on the right-hand side is obtained from  $\mathcal{A}^{(1\text{-loop})}$  in the same way as  $\tilde{L}^{(N)}$  is obtained from  $L^{(N)}$ : starting from any Feynman diagram in  $\mathcal{A}^{(1\text{-loop})}$ , and considering

all possible replacements of each Feynman propagator  $G(q_i)$  in the loop with the cut propagator  $\tilde{\delta}(q_i; M_i)$ , and then replacing the uncut Feynman propagators with dual propagators. All the other factors in the Feynman diagrams are left unchanged in going from  $\mathcal{A}^{(1\text{-loop})}$  to  $\tilde{\mathcal{A}}^{(1\text{-loop})}$ .

Equation (22) establishes a correspondence between the one-loop Feynman diagrams contributing to  $\mathcal{A}^{(1\text{-loop})}$  and the tree-level Feynman diagrams contributing to the phase-space integral in  $\tilde{\mathcal{A}}^{(1\text{-loop})}$ . How are these tree-level Feynman diagrams related to those contributing to the tree-level expression  $\mathcal{A}^{(\text{tree})}$ , i.e. the tree-level counterpart of  $\mathcal{A}^{(1\text{-loop})}$ ? The answer to this question is mainly a matter of combinatorics of Feynman diagrams. If  $\mathcal{A}^{(1\text{-loop})}$  is an off-shell Green's function, the phase-space integrand in  $\tilde{\mathcal{A}}^{(1\text{-loop})}$  is directly related to  $\mathcal{A}^{(\text{tree})}$  [32]. In a sketchy form, we can write:

$$\mathcal{A}_N^{(1\text{-loop})}(\dots) \sim \int_q \sum_P \tilde{\delta}(q; M_P) \tilde{\mathcal{A}}_{N+2}^{(\text{tree})}(q, -q, \dots), \quad (23)$$

where  $\sum_P$  denotes the sum over all the types of particles and antiparticles that can propagate in the loop internal lines, and  $\tilde{\mathcal{A}}^{(\text{tree})}$  simply differs from  $\mathcal{A}^{(\text{tree})}$  by the replacement of dual and Feynman propagators. The extension of Eq. (23) to scattering amplitudes requires a careful treatment of the on-shell limit of the corresponding Green's functions [32].

In recent years much progress [20, 34–40] has been achieved on the computation of tree-level amplitudes, including results in compact analytic form. Using the duality relation, this amount of information at the tree level can be exploited for applications to analytic calculations at the one-loop level.

The computation of cross sections at next-to-leading order (NLO) requires the separate evaluation of real and virtual radiative corrections. Real (virtual) radiative corrections are given by multi-leg tree-level (one-loop) matrix elements to be integrated over the multiparticle phase-space of the physical process. The loop–tree duality discussed here, as well as other methods that relate one-loop and phase-space integrals, have an attractive feature [30, 41–44]: they recast the virtual radiative corrections in a form that closely parallels the contribution of the real radiative corrections. This close correspondence can help to directly combine real and virtual contributions to NLO cross sections. In particular, using the duality relation, we can apply [30] mixed analytical/numerical techniques to the evaluation of the one-loop virtual contributions. The (infrared or ultraviolet) divergent part of the corresponding dual integrals can be analytically evaluated in dimensional regularization. The finite part of the dual integrals can be computed numerically, together with the finite part of the real emission contribution. Partial results along these lines are presented in Refs. [30, 31] and further work is in progress. The extension of the duality relation from one-loop to two-loop Feynman diagrams is also under investigation.

### 3 Monte Carlo simulations of $t\bar{t}$ + jets at hadron colliders

*Author: Michele Treccani*

Because of the high energy of the Tevatron and the LHC, one of the most interesting fields refers to the class of events with multiple final states, giving rise to multiple jets with complicated topologies. There exist different strategies to tackle this problem, with distinct features and points of strength. The main problem is how to consistently compose the contributions due to Matrix Element (ME) calculations with the contributions of the Monte Carlo (MC) showering codes, in

order to exploit their complementarity and avoid at the same time the so-called double counting phenomenon [45–48].

We will here focus on a particular approach which relies on a consistent leading-logarithmic (LL) accuracy in the prediction of a final state  $F$  accompanied by a varying number of extra jets. The double counting is avoided adopting a so-called *matching algorithm* for matrix elements and parton shower. We study in detail the *MLM* matching [49–51] embedded in the the ME generator ALPGEN [52] in order to describe the  $t\bar{t}$  pair production at hadron colliders. First we will address its stability with respect to its internal parameters by comparing predictions obtained with different parameters.

In a step further, we will perform detailed numerical comparison between *MLM* matching and MC program MC@NLO which is an alternative strategy to cope with double counting and reaches next-to-leading order (NLO) accuracy in the prediction [53–55].

### 3.1 Consistency studies of the matching algorithm

In this section we study the overall consistency of the matching algorithm applied to the case of  $t\bar{t}$  final states. We shall consider  $t\bar{t}$  production at the Tevatron ( $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV) and at the LHC ( $pp$  collisions at  $\sqrt{s} = 14$  TeV).

The generation parameters for the light partons are defined by the following kinematical cuts: the default values for the event samples at the Tevatron (LHC) are given by:  $p_T^{min}=20$  (30) GeV and  $R_{min}=0.7$  (0.7), while they are considered only in the geometrical region defined by  $\eta \leq 4$  (5).

The top particle is assumed to be stable, and therefore all jets coming from the decay of top quarks are neglected. For the shower evolution we use HERWIG, version 6.510 [56–58]. We stopped the evolution after the perturbative phase, in order to drop down all the common systematics that could smooth out any possible discrepancy between the various simulations. For all generations we chose the parton distribution function set MRST2001J [59], with renormalization and factorization scales squared set equal to:

$$\mu_R^2 = \mu_F^2 = \sum_{i=t,\bar{t},\text{jets}} [m_i^2 + (p_T^i)^2].$$

Jet observables are built out of the partons emerging from the shower in the rapidity range  $|\eta| \leq 6$  and adopting the cone algorithm GETJET [60]. The jet cone size is set to  $R_{cone} = 0.7$  and the minimum transverse momentum to define a jet at the Tevatron(LHC) is 15(20) GeV .

To our analysis, the important feature of the whole procedure is the presence of two set of parameters: the generation cuts and the matching cuts (see [49–51]). The first set is necessary to avoid the infrared (IR) and collinear singularities:  $p_T^{min}$ , the minimum transverse momentum of the extra parton(s) to be generated, and  $R_{min}$ , the minimum separation between extra-partons in the  $(\eta, \phi)$  plane. Along with these parameters, there exist an analogous set, but with slightly different meanings : the matching cuts  $E_T^{clus}$  and  $R_{match}$ .

We choose two independent variations of the generation and of two of the matching cuts, while keeping fixed our definition of the physical objects (the jets) and of the observables. In both cases, we find that these distribution are stable against reasonable variations of the internal parameters, with relative differences confined well below few percents.

Angular observables, such as  $\Delta R$  between jets, are more sensible, since they are directly related to the matching variables, nevertheless their agreement is within 10%.

The analysis at the LHC, which will not be shown here, leads to qualitatively and quantitatively similar results.

### 3.2 Comparisons with MC@NLO

We shall now compare in detail the description of  $t\bar{t}$  events as provided by ALPGEN and MC@NLO. For consistency with the MC@NLO approach, where only the  $\mathcal{O}(\alpha_s^3)$  ME effects are included, we use ALPGEN samples obtained by stopping the ME contributions only to 1 extra-parton besides the  $t\bar{t}$  pair. This strategy allow to highlight the different features of the two alternative approaches applied to same set of contributions. It is understood that a homogeneous comparison can only be done through the introduction of a proper K-factor, determined by the ratio of the total rates of the two predictions. We adopt the same simulation setup as before, modifying only the same factorization and renormalization scale in order to match MC@NLO's default:

$$\mu_R^2 = \mu_F^2 = \sum_{i=t,\bar{t}} \frac{1}{2} [m_i^2 + (p_T^i)^2].$$

The upper two rows of plots in Fig. 3 refer to inclusive properties of the  $t\bar{t}$  system, namely the transverse momentum and rapidity of the top and anti-top quark, the transverse momentum of the  $t\bar{t}$  pair, and the azimuthal angle  $\Delta\phi^{t\bar{t}}$  between the top and anti-top quark. The overall agreement is good, once ALPGEN is corrected with the proper K-factor (1.36 for the Tevatron, and 1.51 for the LHC), and no large discrepancy is seen between the two descriptions of the chosen distributions. The most significant differences (10 to 20%) are seen in the  $p_T^{top}$  distribution, ALPGEN's one being slightly softer.

In jet-related quantities, while the  $p_T$  of leading and sub-leading jets agree, instead the rapidity of the leading jet reveals two distinct patterns: MC@NLO predictions show a dip at  $y_1 = 0$ , which is not present in ALPGEN predictions. This difference is particularly marked at the Tevatron, but is very visible also at the LHC. This is shown in the right figure of the third row in Fig. 3. Visible differences are also present in the distribution of the first and second jet separation in  $(\eta, \phi)$  space,  $\Delta R_{1,2}$ . To understand the difference in the rapidity distribution, we look in more detail in Fig. 4 at some features in the MC@NLO description of the leading jet. For the  $p_T$  of the leading jet,  $p_{T,1}$ , we plot separately the contribution from the various components of the MC@NLO generation: events in which the shower is initiated by the LO  $t\bar{t}$  hard process, and events in which the shower is initiated by a  $t\bar{t} + q(g)$  hard process. In the latter we separate the contribution of positive- and negative-weight events, where the distribution of negative events is shown in absolute value. The plots show that for MC@NLO the contribution of the  $t\bar{t} + q(g)$  hard process is almost negligible over most of the relevant range and becomes appreciable only for very large values of  $p_{T,1}$ . This hierarchy is stronger at the LHC than at the Tevatron.

Upper set of Fig. 5 shows the various contributions to the rapidity distribution  $y_1$  for different jet  $p_T$  thresholds. It appears that the  $y_1$  distribution resulting from the shower evolution of the  $t\bar{t}$  events in MC@NLO has a strong dip at  $y_1=0$ , a dip that cannot be compensated by the more central distributions of the jet from the  $t\bar{t} + q(g)$  hard process, given its marginal role in the overall jet rate.

That the dip at  $y_1=0$  is a feature typical of jet emission from the  $t\bar{t}$  state in HERWIG is shown in central set of Fig. 5, obtained from the standard HERWIG code rather than from MC@NLO. We speculate that this feature is a consequence of the dead-cone description of hard emission from heavy quarks implemented in the HERWIG shower algorithm. To complete our analysis, we show in lower set of Fig. 5 the comparison between the ALPGEN, MC@NLO and the parton-level  $y_1$  spectra, for different jet  $p_T$  thresholds. We notice that at large  $p_T$ , where the Sudakov effects that induce potential differences between the shower and the PL results have vanished, the ALP-

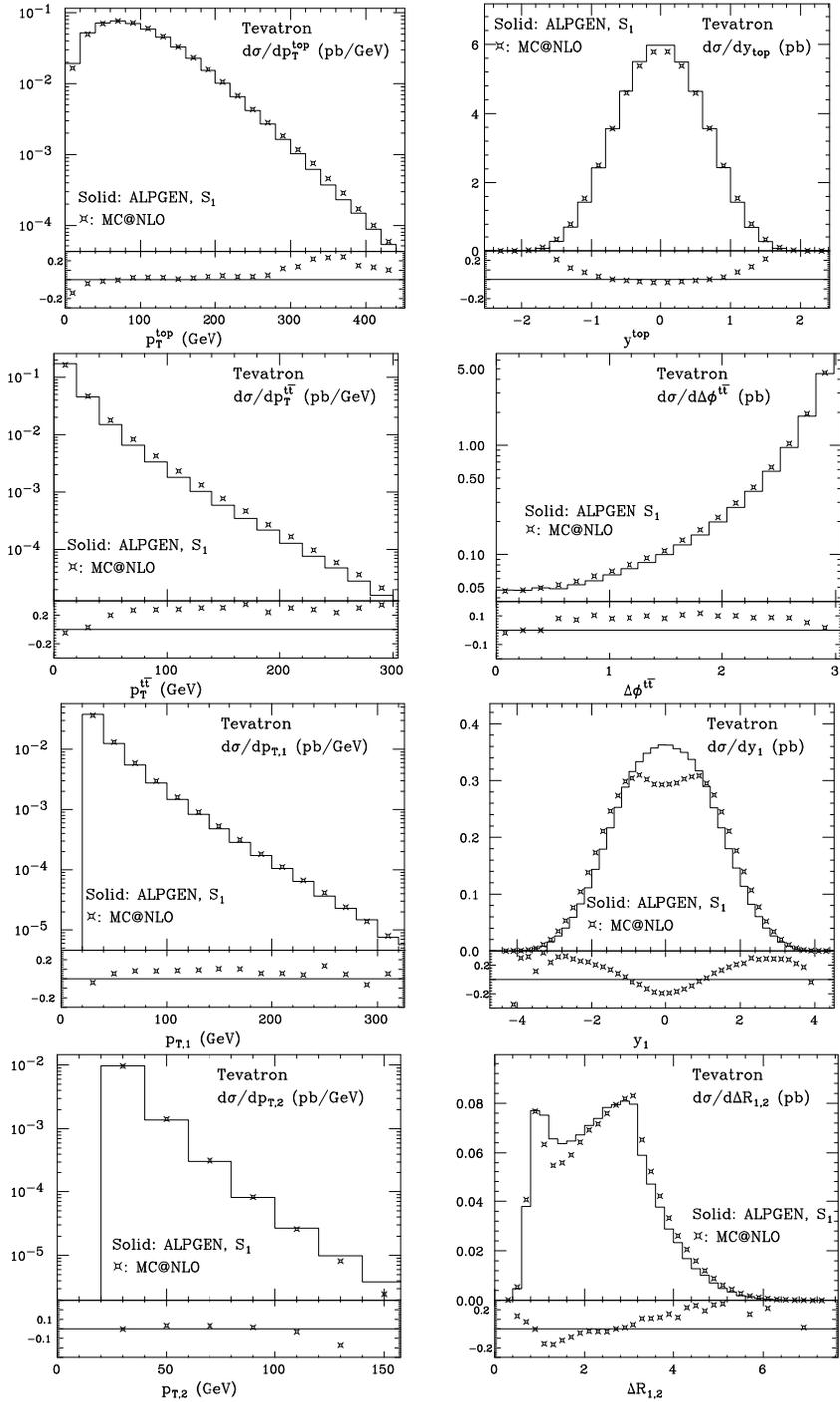


Fig. 3: Comparison of ALPGEN (histogram) and MC@NLO (plot) distributions, at the Tevatron. The ALPGEN results are rescaled to MC@NLO, using the K factor of 1.36. The relative difference  $(\text{MC@NLO} - \text{ALPGEN})/\text{ALPGEN}$  is shown at the bottom of each plot.

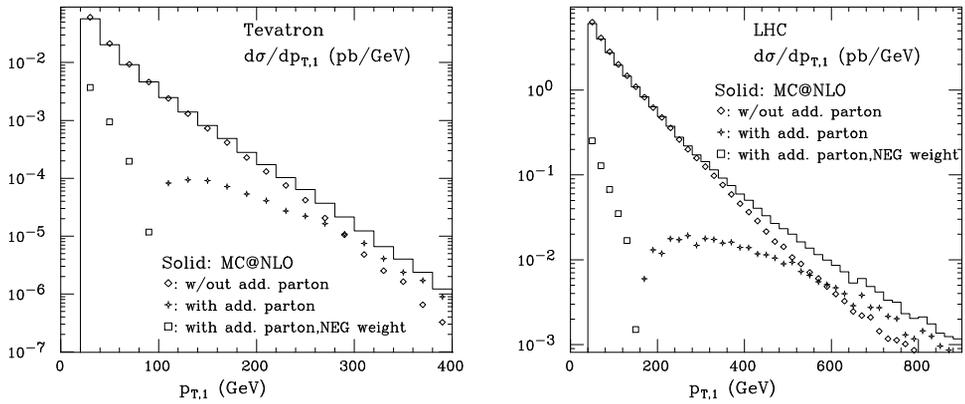


Fig. 4: Contributions to the transverse momentum of the leading jet in MC@NLO. Tevatron (left) and LHC (right).

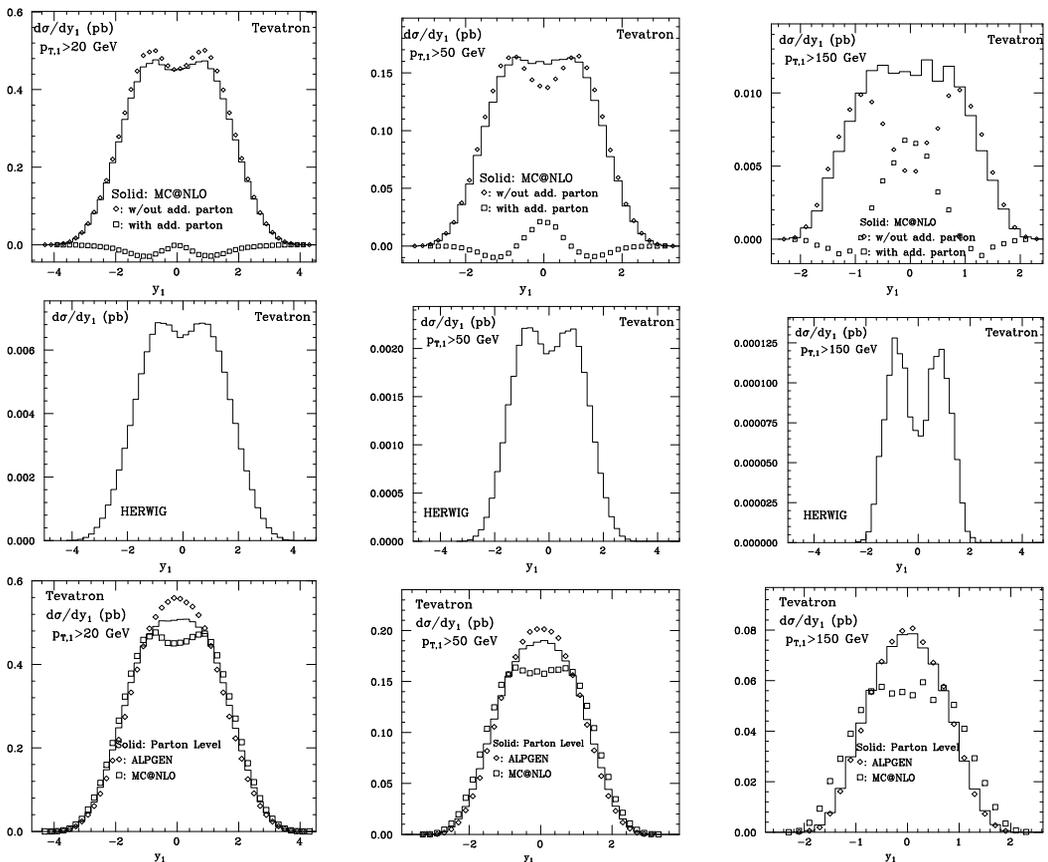


Fig. 5: Rapidity of the leading jet  $y_1$  at Tevatron for various jet  $p_T$  thresholds. Upper set: MC@NLO, with partial contributions. Central set: HERWIG. Lower set: comparison between ALPGEN, MC@NLO, and the parton level predictions

GEN result reproduces well the PL result, while still differing significantly from the MC@NLO distributions.

### 3.3 Conclusions

The analysis presented here is focused on the MC simulations of the  $t\bar{t}$ +jets process as predicted by ALPGEN and its matching algorithm. Several checks of that algorithm have shown its internal consistency, and pinpoint a mild dependence of the results on the parameters that define it. The consistency of the approach is then confirmed by the comparison with MC@NLO. In particular, inclusive variables show excellent agreement, once the NLO/LO K-factor is included.

Instead we found a rather surprising difference between the predictions of two codes for the rapidity distribution of the leading jet accompanying the  $t\bar{t}$  pair. In view of the relevance of this variable for the study at the LHC of new physics signals, it is important to further investigate the origin of this discrepancy, with independent calculations, and with a direct comparison with data. Preliminary results obtained with the new positive-weight NLO shower MC introduced in [61–63] appear to support the distributions predicted by ALPGEN.

## 4 A subtraction scheme for jet cross sections at NNLO

*Author: Gabor Somogyi*

One of the main difficulties in performing NNLO calculations is that the finite higher-order corrections are sums of several pieces which are separately infrared (IR) divergent in  $d = 4$  spacetime dimensions. To handle the IR singularities present in the intermediate stages of calculation in a general (process- and observable-independent) way is non-trivial already at NLO accuracy, where however several solutions are known [2, 3, 64–67]. It is perhaps fair to say that the most widely used is the dipole subtraction scheme of Ref. [2], which constructs a completely general and fully local approximate cross section to regularize real radiation at NLO. Setting up a general subtraction algorithm analogous to that of Ref. [2] but at NNLO accuracy has proved to be rather difficult problem. Here we give a progress report on constructing such a scheme.

### 4.1 Subtraction scheme at NNLO

In perturbative QCD the formal loop expansion for any production rate to NNLO accuracy reads

$$\sigma = \sigma^{\text{LO}} + \sigma^{\text{NLO}} + \sigma^{\text{NNLO}} + \dots \quad (24)$$

Let us consider  $e^+e^- \rightarrow m$  jet production. Then the NNLO correction may be written as

$$\sigma^{\text{NNLO}} = \int_{m+2} d\sigma_{m+2}^{\text{RR}} J_{m+2} + \int_{m+1} d\sigma_{m+1}^{\text{RV}} J_{m+1} + \int_m d\sigma_m^{\text{VV}} J_m, \quad (25)$$

*i.e.* it is the sum of a doubly-real, a real-virtual and a doubly-virtual contribution, each IR divergent in  $d = 4$  spacetime dimensions.

The general strategy of subtraction consists of the following steps: (i) we regularize all integrals in Eq. (25) by dimensional regularization then (ii) we reshuffle the singularities between the three terms by adding and subtracting suitably defined *approximate cross sections* so that finally we rewrite Eq. (25) as

$$\sigma^{\text{NNLO}} = \int_{m+2} d\sigma_{m+2}^{\text{NNLO}} + \int_{m+1} d\sigma_{m+1}^{\text{NNLO}} + \int_m d\sigma_m^{\text{NNLO}}, \quad (26)$$

where now each term on the right hand side is finite in  $d = 4$  by construction. According to Ref. [68] we have

$$d\sigma_{m+2}^{\text{NNLO}} = \left\{ d\sigma_{m+2}^{\text{RR}} J_{m+2} - d\sigma_{m+2}^{\text{RR},A_2} J_m - \left[ d\sigma_{m+2}^{\text{RR},A_1} J_{m+1} - d\sigma_{m+2}^{\text{RR},A_{12}} J_m \right] \right\}_{\varepsilon=0}, \quad (27)$$

$$d\sigma_{m+1}^{\text{NNLO}} = \left\{ \left[ d\sigma_{m+1}^{\text{RV}} + \int_1 d\sigma_{m+2}^{\text{RR},A_1} \right] J_{m+1} - \left[ d\sigma_{m+1}^{\text{RV},A_1} + \left( \int_1 d\sigma_{m+2}^{\text{RR},A_1} \right)^{A_1} \right] J_m \right\}_{\varepsilon=0} \quad (28)$$

and

$$d\sigma_m^{\text{NNLO}} = \left\{ d\sigma_m^{\text{VV}} + \int_2 \left[ d\sigma_{m+2}^{\text{RR},A_2} - d\sigma_{m+2}^{\text{RR},A_{12}} \right] + \int_1 \left[ d\sigma_{m+1}^{\text{RV},A_1} + \left( \int_1 d\sigma_{m+2}^{\text{RR},A_1} \right)^{A_1} \right] \right\}_{\varepsilon=0} J_m. \quad (29)$$

In Eq. (27) above  $d\sigma_{m+2}^{\text{RR},A_1}$  and  $d\sigma_{m+2}^{\text{RR},A_2}$  regularize the singly- and doubly-unresolved limits of  $d\sigma_{m+2}^{\text{RR}}$  respectively. The role of  $d\sigma_{m+2}^{\text{RR},A_{12}}$  is two-fold: it must regularize the singly-unresolved limits of  $d\sigma_{m+2}^{\text{RR},A_2}$  and the doubly-unresolved limits of  $d\sigma_{m+2}^{\text{RR},A_1}$  *simultaneously*. In Eq. (28)  $d\sigma_{m+1}^{\text{RV},A_1}$  and  $\left( \int_1 d\sigma_{m+2}^{\text{RR},A_1} \right)^{A_1}$  regularize the singly-unresolved limits of  $d\sigma_{m+1}^{\text{RV}}$  and  $\int_1 d\sigma_{m+2}^{\text{RR},A_1}$  respectively.

## 4.2 Devising approximate cross sections

Attempting to use the known (multiple) IR factorization properties of (one-loop) squared matrix elements to devise the approximate cross sections in Eqs. (27) and (28) above, we are immediately faced with two problems. First, the various limits overlap in some regions of phase space, thus care needs to be taken to avoid multiple subtraction. Second, even once the factorization formulae are written in such a way that intersecting limits are disentangled so that multiple subtraction does not occur, the resulting expressions cannot be used as true subtraction terms because they are only defined in the strict soft and/or collinear limits. Thus, constructing the approximate cross sections proceeds in two steps: (i) we write all relevant factorization formulae in such a way that their overlap structure can be disentangled (“*matching of limits*”) and (ii) we define “*extensions*” of the formulae so that they are unambiguously defined away from the IR limits.

Let us consider first the matching of limits. A single parton, say  $r$ , can become unresolved in (i) the collinear limit, when for some hard parton  $i \neq r$  we have  $p_i \parallel p_r$  and (ii) in the soft limit, when  $p_r \rightarrow 0$ . In these limits QCD squared matrix elements obey well-known universal factorization properties [69–72], which we exhibit below at tree level for the sake of being specific<sup>1</sup>

$$\mathbf{C}_{ir} |\mathcal{M}_{m+2}^{(0)}|^2 \propto \frac{1}{s_{ir}} \langle \mathcal{M}_{m+1}^{(0)} | \hat{P}_{ir}^{(0)}(z_i, k_\perp; \varepsilon) | \mathcal{M}_{m+1}^{(0)} \rangle, \quad (30)$$

$$\mathbf{S}_r |\mathcal{M}_{m+2}^{(0)}|^2 \propto \sum_{\substack{i,k \\ i \neq k}} \frac{s_{ik}}{s_{ir} s_{kr}} \langle \mathcal{M}_{m+1}^{(0)} | \mathbf{T}_i \mathbf{T}_r | \mathcal{M}_{m+1}^{(0)} \rangle. \quad (31)$$

To write Eqs. (30) and (31) above, we used the colour-state notation of Ref. [2] and the operator notation of taking the limits introduced in Ref. [73], while  $s_{jl} = 2p_j \cdot p_l$ , ( $j, l = i, k, r$ ),  $\hat{P}_{ir}^{(0)}$

<sup>1</sup>To keep the discussion as simple as possible, we only indicate the structure of the factorization formulae.

are the tree-level Altarelli–Parisi splitting kernels and finally  $z_i$  is the momentum-fraction carried by parton  $i$  in the  $p_{ir} \rightarrow p_i + p_r$  splitting. When parton  $r$  is both soft and collinear to the hard parton  $i$ , these limits overlap. To avoid double subtraction in this region of phase space, we must identify the common soft-collinear limit of Eqs. (30) and (31), which is found to be [73]

$$\mathbf{C}_{ir} \mathbf{S}_r |\mathcal{M}_{m+2}^{(0)}|^2 \propto \frac{1}{s_{ir}} \frac{2z_i}{1-z_i} \mathbf{T}_i^2 |\mathcal{M}_{m+1}^{(0)}|^2. \quad (32)$$

Thus the formal operator

$$\mathbf{A}_1 = \sum_r \left[ \sum_{i \neq r} \frac{1}{2} \mathbf{C}_{ir} + \left( \mathbf{S}_r - \sum_{i \neq r} \mathbf{C}_{ir} \mathbf{S}_r \right) \right] \quad (33)$$

counts each singly-unresolved limit precisely once and is free of double subtractions, therefore  $\mathbf{A}_1 |\mathcal{M}_{m+2}^{(0)}|^2$  has the same singly-unresolved singularity structure as  $|\mathcal{M}_{m+2}^{(0)}|^2$  itself, *i.e.* it defines a candidate subtraction term for constructing  $d\sigma_{m+2}^{\text{RR}, \mathbf{A}_1}$ . Similarly, applying the formal operator  $\mathbf{A}_1$  to *e.g.*  $2\Re \langle \mathcal{M}_{m+1}^{(0)} | | \mathcal{M}_{m+1}^{(1)} \rangle$  defines a candidate subtraction term for defining  $d\sigma_{m+1}^{\text{RV}, \mathbf{A}_1}$ , starting from the collinear [11, 74–76] and soft [77] factorization formulae for one-loop squared matrix elements.

The matching procedure is quite a bit more elaborate when two different partons, say  $r$  and  $s$ , become unresolved, which can arise in four different limits: (i) the triple collinear limit, when for some hard parton  $i \neq r, s$  we have  $p_i || p_r || p_s$ , (ii) the doubly single collinear limit, when for two distinct hard partons  $i \neq r, s$  and  $j \neq r, s$  we have  $p_i || p_r$  and  $p_j || p_s$ , (iii) the doubly soft-collinear limit, when for  $i \neq r, s$  we have  $p_i || p_r$  and  $p_s \rightarrow 0$ , and finally (iv) the double soft limit, when  $p_r \rightarrow 0$  and  $p_s \rightarrow 0$ . The factorization formulae appropriate for each of these limits are well-known (in particular the three-parton splitting functions and the double soft  $gg$  and  $q\bar{q}$  currents are given in Refs. [78–84] and Refs. [72, 85], respectively), and their highly non-trivial overlap structure was disentangled in Ref. [73]. To identify the intersection of limits, Ref. [73] computed all common limits explicitly, which is rather cumbersome. In [86], a simple and systematic procedure was proposed that leads directly to pure soft factorization formulae at any order and thus solves the problem of matching of limits in general. Finally (using the operator notation of Ref. [73]) we find that the symbolic operator

$$\begin{aligned} \mathbf{A}_2 = & \sum_r \sum_{s \neq r} \left\{ \sum_{i \neq r, s} \left[ \frac{1}{6} \mathbf{C}_{irs} + \sum_{j \neq i, r, s} \frac{1}{8} \mathbf{C}_{ir;js} + \frac{1}{2} \mathbf{C} \mathbf{S}_{ir;s} \right] + \frac{1}{2} \mathbf{S}_{rs} - \sum_{i \neq r, s} \left[ \frac{1}{2} \mathbf{C}_{irs} \mathbf{C} \mathbf{S}_{ir;s} \right. \right. \\ & \left. \left. + \sum_{j \neq i, r, s} \frac{1}{2} \mathbf{C}_{ir;js} \mathbf{C} \mathbf{S}_{ir;s} + \frac{1}{2} \mathbf{C}_{irs} \mathbf{S}_{rs} + \mathbf{C} \mathbf{S}_{ir;s} \mathbf{S}_{rs} - \sum_{j \neq i, r, s} \frac{1}{2} \mathbf{C}_{ir;js} \mathbf{S}_{rs} - \mathbf{C}_{irs} \mathbf{C} \mathbf{S}_{ir;s} \mathbf{S}_{rs} \right] \right\} \quad (34) \end{aligned}$$

counts each doubly-unresolved limit precisely once (without overlaps). Thus  $\mathbf{A}_2 |\mathcal{M}_{m+2}^{(0)}|^2$  has the same doubly-unresolved singularity structure as  $|\mathcal{M}_{m+2}^{(0)}|^2$  itself and so defines a candidate subtraction term for constructing  $d\sigma_{m+2}^{\text{RR}, \mathbf{A}_2}$ .

Finally, we must address the matching of the singly- and doubly-unresolved limits of  $|\mathcal{M}_{m+2}^{(0)}|^2$  which also overlap.  $d\sigma_{m+2}^{\text{RR}, \mathbf{A}_{12}}$  is introduced in Eq. (27) precisely to avoid double subtraction in the intersecting regions of phase space. However the role of this approximate cross

section is quite delicate, because (i) in the doubly-unresolved limits it must regularize  $d\sigma_{m+1}^{\text{RR},A_1}$ , while (ii) in the singly-unresolved limits, it must regularize  $d\sigma_{m+2}^{\text{RR},A_2}$  and spurious singularities that appear in  $d\sigma_{m+2}^{\text{RR},A_1}$ . It is thus a highly non-trivial statement that the correct candidate subtraction term can be obtained by applying the symbolic singly-unresolved operator  $\mathbf{A}_1$  of Eq. (33) to  $\mathbf{A}_2|\mathcal{M}_{m+2}^{(0)}|^2$  [73]. That is,

$$(\mathbf{A}_1 + \mathbf{A}_2 - \mathbf{A}_1\mathbf{A}_2)|\mathcal{M}_{m+2}^{(0)}|^2 \quad (35)$$

has the same singularity structure as  $|\mathcal{M}_{m+2}^{(0)}|^2$  itself in all singly- and doubly-unresolved limits and is free of multiple subtractions.

The second step of defining the approximate cross sections calls for an extension of the limit formulae over the full phase space. As emphasized above, the candidate subtraction terms cannot yet be used as true subtraction terms because they are only well-defined in the strict limits. In order to define suitable extensions over the full phase space, we need to define momentum mappings  $\{p\}_{m+2} \rightarrow \{\tilde{p}\}_{m+1}$  and  $\{p\}_{m+2} \rightarrow \{\tilde{p}\}_m$  that (i) implement exact momentum conservation, (ii) lead to exact phase space factorization and (iii) respect the delicate structure of cancellations among the subtraction terms in the various limits. We find it convenient to define two types of singly-unresolved ( $\{p\}_{m+2} \rightarrow \{\tilde{p}\}_{m+1}$ ) mappings and four types of doubly-unresolved ( $\{p\}_{m+2} \rightarrow \{\tilde{p}\}_m$ ) mappings, corresponding to the basic types of limits that may occur (*i.e.* we define a collinear and a soft singly-unresolved mapping). The explicit forms of these momentum mappings may be found in Ref. [68] together with the full definitions of all approximate cross sections that appear in Eq. (27). The approximate cross sections in Eq. (28) are given explicitly in Refs. [87, 88].

At the risk of belabouring the point, we note again that all our momentum mappings lead to an exact factorization of the phase space in the symbolic form

$$d\phi_{m+2} = d\phi_{m+1}[dp_1] \quad \text{and} \quad d\phi_{m+2} = d\phi_m[dp_2], \quad (36)$$

thus the singular integrals of the subtraction terms over the phase space of the unresolved parton(s) can be computed once and for all, independent of the jet function and the rest of the phase space integration.

### 4.3 Conclusions

We have set up a general (process- and observable-independent) subtraction scheme for computing QCD jet cross sections at NNLO accuracy for processes with no coloured particles in the initial state. Our scheme can naturally be viewed as the generalization of the dipole subtraction scheme of Ref. [2] to NNLO. We have defined all approximate cross sections needed to regularize the  $m+2$  and  $m+1$  parton contributions (*i.e.* all terms in Eqs. (27) and (28)) explicitly. Our subtraction terms are *fully local*, *i.e.* all colour and azimuthal correlations are properly taken into account. Thus we can check the convergence of our subtraction terms to the doubly-real, or real-virtual cross sections in any unresolved limit explicitly. In addition, we have checked that the regularized doubly-real and real-virtual contributions to  $e^+e^- \rightarrow 3$  jet production are finite by computing the first three moments of the thrust and  $C$ -parameter distributions, see Tab. 1. In order to finish the definition of the subtraction scheme, one must still compute the singly- and doubly-unresolved integrals of the approximate cross sections that appear in Eq. (29). All

n	$\langle(1-t)^n\rangle_{\text{RV}}/10^1$	$\langle C^n\rangle_{\text{RV}}/10^1$	$\langle(1-t)^n\rangle_{\text{RR}}$	$\langle C^n\rangle_{\text{RR}}$
1	$123 \pm 1$	$433 \pm 5$	$-92.7 \pm 3.4$	$-344 \pm 14$
2	$25.5 \pm 0.2$	$325 \pm 2$	$-3.07 \pm 0.43$	$-142 \pm 3$
3	$4.79 \pm 0.03$	$180 \pm 1$	$2.01 \pm 0.12$	$6.29 \pm 1.87$

Table 1: The real-virtual and doubly-real contributions to the first three moments of the thrust and  $C$ -parameter distribution in  $e^+e^- \rightarrow 3$  jets.

singly-unresolved integrals (denoted by  $\int_1$  in Eqs. (28) and (29) above) have recently been computed [87, 89–91] and we expect that the techniques applied will be flexible enough to compute the doubly-unresolved integrals (denoted by  $\int_2$  in Eq. (29)) as well. This is work in progress.

We are grateful to our collaborators: U. Aglietti, P. Bolzoni, V. Del Duca, C. Duhr and S. Moch. This work was supported in part by Hungarian Scientific Research Fund grant OTKA K-60432 and by the Swiss National Science Foundation (SNF) under contract 200020-117602.

## References

- [1] R. K. Ellis, D. A. Ross, and A. E. Terrano, Nucl. Phys. **B178**, 421 (1981).
- [2] S. Catani and M. H. Seymour, Nucl. Phys. **B485**, 291 (1997), [arXiv:hep-ph/9605323](#).
- [3] S. Frixione, Z. Kunszt, and A. Signer, Nucl. Phys. **B467**, 399 (1996), [arXiv:hep-ph/9512328](#).
- [4] E. W. N. Glover, Nucl. Phys. Proc. Suppl. **116**, 3 (2003), [arXiv:hep-ph/0211412](#).
- [5] W. T. Giele and E. W. N. Glover, JHEP **04**, 029 (2004), [arXiv:hep-ph/0402152](#).
- [6] R. K. Ellis, W. T. Giele, and G. Zanderighi, Phys. Rev. **D73**, 014027 (2006), [arXiv:hep-ph/0508308](#).
- [7] A. Denner and S. Dittmaier, Nucl. Phys. **B734**, 62 (2006), [arXiv:hep-ph/0509141](#).
- [8] T. Binoth, J. P. Guillet, G. Heinrich, E. Pilon, and T. Reiter (2008), [arXiv:0810.0992 \[hep-ph\]](#).
- [9] W. Giele, E. W. N. Glover, and G. Zanderighi, Nucl. Phys. Proc. Suppl. **135**, 275 (2004), [arXiv:hep-ph/0407016](#).
- [10] NLO Multileg Working Group Collaboration, Z. Bern *et al.* (2008), [arXiv:0803.0494 \[hep-ph\]](#).
- [11] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, Nucl. Phys. **B425**, 217 (1994), [arXiv:hep-ph/9403226](#).

- [12] R. Britto, F. Cachazo, and B. Feng, Nucl. Phys. **B725**, 275 (2005), arXiv:hep-th/0412103.
- [13] C. F. Berger, Z. Bern, L. J. Dixon, D. Forde, and D. A. Kosower, Phys. Rev. **D75**, 016006 (2007), arXiv:hep-ph/0607014.
- [14] C. F. Berger *et al.*, Phys. Rev. **D78**, 036003 (2008), arXiv:0803.4180 [hep-ph].
- [15] R. K. Ellis, W. T. Giele, and Z. Kunszt, JHEP **03**, 003 (2008), arXiv:0708.2398 [hep-ph].
- [16] W. T. Giele, Z. Kunszt, and K. Melnikov, JHEP **04**, 049 (2008), arXiv:0801.2237 [hep-ph].
- [17] W. T. Giele and G. Zanderighi, JHEP **06**, 038 (2008), arXiv:0805.2152 [hep-ph].
- [18] G. Ossola, C. G. Papadopoulos, and R. Pittau, Nucl. Phys. **B763**, 147 (2007), arXiv:hep-ph/0609007.
- [19] Z. Bern and D. A. Kosower, Nucl. Phys. **B362**, 389 (1991).
- [20] F. Berends and W. Giele, Nucl. Phys. **B306**, 759 (1988).
- [21] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, Nucl. Phys. **B435**, 59 (1995), arXiv:hep-ph/9409265.
- [22] R. K. Ellis and G. Zanderighi, JHEP **02**, 002 (2008), arXiv:0712.1851 [hep-ph].
- [23] R. K. Ellis, W. T. Giele, and G. Zanderighi, JHEP **05**, 027 (2006), arXiv:hep-ph/0602185.
- [24] D. Forde and D. A. Kosower, Phys. Rev. **D73**, 061701 (2006), arXiv:hep-ph/0509358.
- [25] R. Kleiss, W. J. Stirling, and S. D. Ellis, Comput. Phys. Commun. **40**, 359 (1986).
- [26] V. Del Duca, W. Kilgore, C. Oleari, C. Schmidt, and D. Zeppenfeld, Nucl. Phys. **B616**, 367 (2001), arXiv:hep-ph/0108030.
- [27] R. Kleiss and H. Kuijf, Nucl. Phys. **B312**, 616 (1989).
- [28] R. K. Ellis, W. T. Giele, Z. Kunszt, and K. Melnikov (2008), arXiv:0806.3467 [hep-ph].
- [29] R. K. Ellis, W. T. Giele, Z. Kunszt, K. Melnikov, and G. Zanderighi (2008), arXiv:0810.2762 [hep-ph].
- [30] S. Catani. Presented at the Workshop *HP<sup>2</sup>*: High Precision for Hard Processes at the LHC, Sept. 2006, Zurich, Switzerland.
- [31] T. Gleisberg. Ph.D. Thesis, University of Dresden. .

- [32] S. Catani, T. Gleisberg, F. Krauss, G. Rodrigo, and J.-C. Winter, *JHEP* **09**, 065 (2008), [arXiv:0804.3170 \[hep-ph\]](#).
- [33] R. Feynman, *Acta Phys. Polon.* **24**, 697 (1963).
- [34] M. Mangano and S. Parke, *Phys. Rept.* **200**, 301 (1991), [arXiv:hep-th/0509223](#).
- [35] F. Caravaglios and M. Moretti, *Phys. Lett.* **B358**, 332 (1995), [arXiv:hep-ph/9507237](#).
- [36] R. K. P. Draggiotis and C. Papadopoulos, *Phys. Lett.* **B439**, 157 (1998), [arXiv:hep-ph/9807207](#).
- [37] R. K. P. Draggiotis and C. Papadopoulos, *Eur. Phys. J.* **C24**, 447 (2002), [arXiv:hep-ph/0202201](#).
- [38] P. S. F. Cachazo and E. Witten, *JHEP* **09**, 006 (2004), [arXiv:hep-th/0403047](#).
- [39] F. C. R. Britto and B. Feng, *Nucl. Phys.* **B715**, 499 (2005), [arXiv:hep-th/0412308](#).
- [40] B. F. R. Britto, F. Cachazo and E. Witten, *Phys. Rev. Lett.* **94**, 181602 (2005), [arXiv:hep-th/0501052](#).
- [41] D. Soper, *Phys. Rev. Lett.* **81**, 2638 (1998), [arXiv:hep-ph/9804454](#).
- [42] M. Kramer and D. E. Soper, *Phys. Rev.* **D66**, 054017 (2002), [arXiv:hep-ph/0204113](#).
- [43] T. Kleinschmidt. DESY-THESIS-2007-042.
- [44] M. Moretti, F. Piccinini, and A. D. Polosa (2008), [arXiv:0802.4171 \[hep-ph\]](#).
- [45] R. K. S. Catani, F. Krauss and B. Webber, *JHEP* **11**, 063 (2001).
- [46] L. Lonnblad, *JHEP* **05**, 046 (2002).
- [47] F. Krauss, *JHEP* **08**, 015 (2002).
- [48] S. Hoche *et al.* (2006). [hep-ph/0602031](#).
- [49] M. Mangano (2002). [www-cpd.fnal.gov/personal/mrenna/tuning/nov2002/mlm.pdf](#).
- [50] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, *JHEP* **01**, 013 (2007), [arXiv:hep-ph/0611129](#).
- [51] J. Alwall *et al.*, *Eur. Phys. J.* **C53**, 473 (2008).
- [52] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *JHEP* **07**, 001 (2003), [arXiv:hep-ph/0206293](#).
- [53] S. Frixione and B. R. Webber, *JHEP* **06**, 029 (2002), [hep-ph/0204244](#).

- [54] S. Frixione, P. Nason, and B. R. Webber, *JHEP* **08**, 007 (2003), [hep-ph/0305252](#).
- [55] S. Frixione and B. Webber (2006). The MC@NLO 3.2 event generator, [hep-ph/0601192](#).
- [56] G. Marchesini and B. Webber, *Nucl. Phys.* **B310**, 461 (1988).
- [57] G. Marchesini *et al.*, *Comput. Phys. Commun.* **67**, 465 (1992).
- [58] G. Corcella *et al.*, *JHEP* **01**, 010 (2001), [arXiv:hep-ph/0011363](#).
- [59] W. S. A.D. Martin, R.G. Roberts and R. Thorne, *Eur. Phys. J.* **C23**, 73 (2002).
- [60] E. Paige and S. Protopopescu. Physics of the SSC, in Snowmass, 1986, Colorado, edited by R. Donaldson and J. Marx.
- [61] P. Nason, *JHEP* **11**, 040 (2004), [hep-ph/0409146](#).
- [62] P. Nason and G. Ridolfi, *JHEP* **0608**, 077 (2006).
- [63] C. O. S. Alioli, P. Nason and E. Re, *JHEP* **07**, 060 (2008).
- [64] W. T. Giele and E. W. N. Glover, *Phys. Rev.* **D46**, 1980 (1992).
- [65] W. T. Giele, E. W. N. Glover, and D. A. Kosower, *Nucl. Phys.* **B403**, 633 (1993), [arXiv:hep-ph/9302225](#).
- [66] Z. Nagy and Z. Trócsányi, *Nucl. Phys.* **B486**, 189 (1997), [arXiv:hep-ph/9610498](#).
- [67] S. Frixione, *Nucl. Phys.* **B507**, 295 (1997), [arXiv:hep-ph/9706545](#).
- [68] G. Somogyi, Z. Trócsányi, and V. Del Duca, *JHEP* **01**, 070 (2007), [arXiv:hep-ph/0609042](#).
- [69] J. Frenkel and J. C. Taylor, *Nucl. Phys.* **B116**, 185 (1976).
- [70] D. Amati, R. Petronzio, and G. Veneziano, *Nucl. Phys.* **B146**, 29 (1978).
- [71] A. Bassetto, M. Ciafaloni, and G. Marchesini, *Phys. Rept.* **100**, 201 (1983).
- [72] S. Catani and M. Grazzini, *Nucl. Phys.* **B570**, 287 (2000), [arXiv:hep-ph/9908523](#).
- [73] G. Somogyi, Z. Trócsányi, and V. Del Duca, *JHEP* **06**, 024 (2005), [arXiv:hep-ph/0502226](#).
- [74] Z. Bern, V. Del Duca, and C. R. Schmidt, *Phys. Lett.* **B445**, 168 (1998), [arXiv:hep-ph/9810409](#).
- [75] D. A. Kosower and P. Uwer, *Nucl. Phys.* **B563**, 477 (1999), [arXiv:hep-ph/9903515](#).
- [76] Z. Bern, V. Del Duca, W. B. Kilgore, and C. R. Schmidt, *Phys. Rev.* **D60**, 116001 (1999), [arXiv:hep-ph/9903516](#).

- [77] S. Catani and M. Grazzini, Nucl. Phys. **B591**, 435 (2000), arXiv:hep-ph/0007142.
- [78] A. Gehrmann-De Ridder and E. W. N. Glover, Nucl. Phys. **B517**, 269 (1998), arXiv:hep-ph/9707224.
- [79] J. M. Campbell and E. W. N. Glover, Nucl. Phys. **B527**, 264 (1998), arXiv:hep-ph/9710255.
- [80] S. Catani and M. Grazzini, Phys. Lett. **B446**, 143 (1999), arXiv:hep-ph/9810389.
- [81] D. A. Kosower, Nucl. Phys. **B552**, 319 (1999), arXiv:hep-ph/9901201.
- [82] V. Del Duca, A. Frizzo, and F. Maltoni, Nucl. Phys. **B568**, 211 (2000), arXiv:hep-ph/9909464.
- [83] D. A. Kosower, Phys. Rev. **D67**, 116003 (2003), arXiv:hep-ph/0212097.
- [84] D. A. Kosower, Phys. Rev. Lett. **91**, 061602 (2003), arXiv:hep-ph/0301069.
- [85] F. A. Berends and W. T. Giele, Nucl. Phys. **B313**, 595 (1989).
- [86] Z. Nagy, G. Somogyi, and Z. Trócsányi (2007), arXiv:hep-ph/0702273.
- [87] G. Somogyi and Z. Trócsányi, Acta Phys. Chim. Debr. **XL**, 101 (2006), arXiv:hep-ph/0609041.
- [88] G. Somogyi and Z. Trócsányi, JHEP **01**, 052 (2007), arXiv:hep-ph/0609043.
- [89] G. Somogyi and Z. Trócsányi, JHEP **08**, 042 (2008), arXiv:0807.0509 [hep-ph].
- [90] U. Aglietti, V. Del Duca, C. Duhr, G. Somogyi, and Z. Trócsányi (2008), arXiv:0807.0514 [hep-ph].
- [91] P. Bolzoni, S. Moch, G. Somogyi, and i. p. Trócsányi, Z. (in preparation), arXiv:?? [hep-ph].

# Event shapes and resummation

*Andrea Banfi, Gennaro Corcella, Mrinal Dasgupta, S. Joseph Gionata Luisoni, Swapan Majhi, B.F.L. Ward, S.A. Yost*

## 1 Event shapes and resummation

For the sake of reliable measurements at present and future colliders, the use of precise QCD calculations is mandatory. Fixed-order calculations discussed in Sec. [1] are accurate enough to predict inclusive observables, such as total cross sections or widths, whereas more exclusive quantities, such as event-shape distributions, exhibit large logarithmic enhancements, corresponding to soft- or collinear-parton radiation, which need to be resummed to all orders to improve the perturbative prediction. Analytical resummation of soft/collinear-enhanced radiation can be performed following the general method in [2–4]. Such resummations are usually based on the approximation of multiple independent emissions, implying factorization of amplitudes and phase spaces, and resulting in the exponentiation of soft/collinear single-parton radiation.

In the following we describe recent progress in the understanding and development of such resummations, including a critical comparison of analytical resummations with partons shower resummations, a discussion of non-global logarithms and recent extraction of the strong coupling using newly available NLLA+NNLO matched predictions.

## 2 Parton showers and resummations for non-global QCD observables

*Authors: Andrea Banfi, Gennaro Corcella and Mrinal Dasgupta*

Resummation of soft and collinear logarithms are usually based on the approximation of multiple independent emissions, implying factorization of amplitudes and phase spaces, and resulting in the exponentiation of soft/collinear single-parton radiation. In fact, a resummed quantity  $\Sigma(L)$ ,  $L$  being a large logarithm of soft or collinear origin, typically reads:

$$\Sigma(L) = \exp [Lg_1(\alpha_S L) + g_2(\alpha_S L) + \alpha_S g_3(L) + \dots], \quad (1)$$

where  $Lg_1$  resums the double logarithms, *i.e.* both soft and collinear,  $\mathcal{O}(\alpha_S^n L^{n+1})$ , while  $g_2$  resums single logarithms  $\mathcal{O}(\alpha_S^n L^n)$ , either soft or collinear, and so forth. Contributions  $\sim \alpha_S^n L^{n+1}$  and  $\sim \alpha_S^n L^n$  are typically classified as leading- (LL) and next-to-leading (NLL) logarithms. However, as we shall point out later on, if  $g_1$  is zero, the LLs will be the ones contained in  $g_2$ .

As an alternative tool to resum large logarithms, one can employ Monte Carlo generators, such as HERWIG [5] or PYTHIA [6], which implement parton showers in the soft/collinear approximation and include models for hadronization and the underlying event. In particular, the evolution variable for the HERWIG showers is equivalent, for soft emissions, to angular ordering [7, 8], which is a reliable approximation in the large- $N_C$  limit for azimuthally-averaged quantities. PYTHIA traditionally orders its cascades according to the virtuality of the splitting parton, with the possibility to reject non-angular-ordered showers. Lately, a new PYTHIA shower model [9] was released, ordering multiple emissions according to the transverse momentum of the radiated parton with respect to the emitter’s direction. Monte Carlo algorithms are correct

up to the double-logarithmic function  $g_1$  and in some cases they can even account for  $g_2$  (see, *e.g.*, [10] for some discussions on comparing parton showers and resummations).

In the following, we shall discuss the so-called non-global observables and compare the results of resummed calculations, with the possible inclusion of the angular-ordering approximation, with those given by Monte Carlo parton showers.

### 2.0.1 Non-global observables

It was recently found out [11] that for some quantities, called non-global observables, as they are sensitive to radiation in a limited region of the phase space, the independent-emission approximation is not sufficient any longer, even at LL level. As a case study, we consider  $e^+e^-$  annihilation into hadrons at the centre-of-mass energy  $Q$  and study the transverse-energy flow in an angular region  $\Omega$ , a limited region in rapidity  $\eta$  and azimuth  $\phi$ :

$$\Sigma(Q, Q_\Omega) = \frac{1}{\sigma} \int_0^{Q_\Omega} dE_t \frac{d\sigma}{dE_t} ; E_t = \sum_{i \in \Omega} E_{ti}. \quad (2)$$

$\Sigma$  was computed in [12] and reads:

$$\Sigma(Q, Q_\Omega) = \exp(-4C_F A_\Omega t) S(t), \quad (3)$$

with

$$A_\Omega = \int d\eta \frac{d\phi}{2\pi} ; t = \frac{1}{2\pi} \int_{Q_\Omega}^{Q/2} \frac{dk}{k} \alpha_S(k). \quad (4)$$

In Eq. (3), the contribution  $\sim \exp(-4C_F A_\Omega t)$  comes after exponentiating single-gluon radiation from the primary  $q\bar{q}$  pair, which constitutes the Born event, whereas  $S(t)$  includes non-global logarithms, due to correlated parton emission in the  $\Omega$  region. The lowest-order contribution to  $S(t)$  goes as  $\alpha_S^2 S_2 \ln^2(Q/Q_\Omega)$ , with  $S_2 \sim C_A C_F$ .  $S_2$  was calculated exactly, while the function  $S(t)$  was computed at all orders in the LL approximation and in the large- $N_C$  limit, by using the evolution algorithm presented in [11]. We point out that, for an observable like  $\Sigma$ , the function  $g_1$  in Eq. (1) is zero, hence the leading logarithms are just  $\sim \alpha_S^n L^n$ : including the non-global function  $S(t)$  is therefore necessary to fully account for LLs.

As in Ref. [13], we wish to investigate whether implementing angular ordering in the evolution algorithm of [11] still leads to acceptable results for  $\Sigma(Q, Q_\Omega)$  and  $S(t)$ . In Fig. 1 we present the leading-order non-global coefficient,  $-S_2/(C_F C_A)$ , according to the full calculation and the angular-ordering approximation, in case  $\Omega$  is a rapidity slice of width  $\Delta\eta = 2.5$ . We also show the cross section  $\Sigma(t)$  yielded by the full leading-log resummed calculation and in the angular-ordering (AO) approximation. For the sake of comparison, we also present the contribution coming from just exponentiating primary single-parton emission.

From Fig. 1 (left), we learn that for small gap sizes the full and AO results agree, while they start to differ once the gap is increased. In both cases,  $S_2$  saturates for large  $\Delta\eta$ , with the AO result being about 10% lower than the full one. As for  $\Sigma(t)$ , the AO approximation is indeed able to include significant part of the full result, whereas the primary-emission contribution lies far above the two other predictions, thus giving unreliable spectra. Considering, *e.g.*,  $t = 0.15$ , corresponding to  $Q = 100$  GeV and  $Q_\Omega = 1$  GeV, the AO and primary results are 10% and 75% above the full one, respectively. It was also shown in Ref. [13] that the results for the non-global function  $S(t)$  are roughly independent of the size of the rapidity gap.

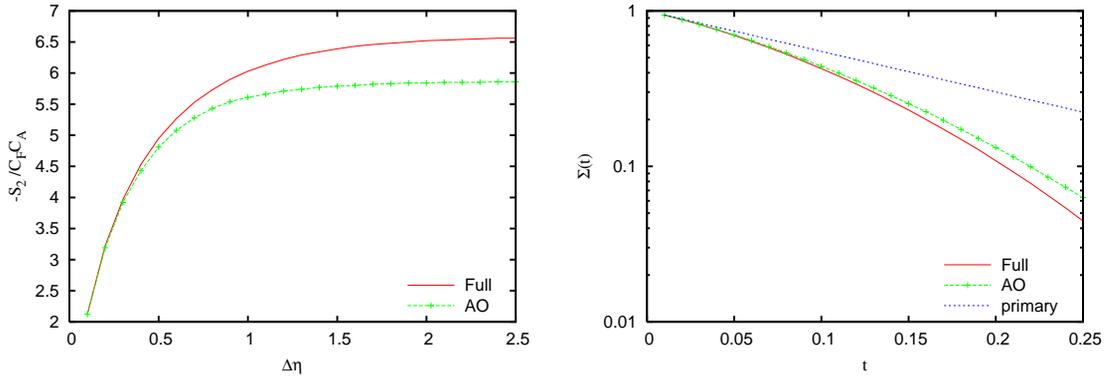


Fig. 1: Left: Function  $S(t)$  at leading-order according to the full LL calculation and in the angular-ordering approximation, in terms of the rapidity gap  $\Delta\eta$ . Right: Function  $\Sigma(t)$  according to the full resummed calculation and the angular-ordering approximation. Also shown is the result coming from the exponentiation of primary-emission contributions.

### 2.0.2 Comparison with HERWIG and PYTHIA

In this section we compare the results of the resummed calculation with the ones yielded by the Monte Carlo programs HERWIG and PYTHIA. As in [13], we study  $e^+e^-$  annihilation at the centre-of-mass energy  $Q = 10^5$  GeV. In fact, we chose such a high value of  $Q$  in order to kill subleading effects, weighted by  $\alpha_S(Q)$  or suppressed by powers of  $1/Q$ , such as subleading soft/collinear logarithms, quark mass effects, hadronization corrections. Furthermore, we checked that our results depend only on the dimensionless variable  $t$  in Eq. (4), so that our findings for a given value of  $t$  can be easily translated to any value of the centre-of-mass energy.

In Fig. 2 we present the differential cross section  $1/\sigma (d\sigma/dE_t)$  for the transverse-energy flow in a rapidity gap  $\Delta\eta = 1$ , according to the resummed result, matched to the exact NLO as in [12], and according to HERWIG and PYTHIA. In the resummation, we show the full result, the angular-ordering approximation and the primary-emission contribution. As for PYTHIA, we present the spectra obtained running the old and new models, with showers ordered in virtuality and transverse momentum, respectively. When using the old model, we shall always assume that non-AO radiation is vetoed.

As for the comparison with HERWIG, whose showers are ordered in angle, we observe good agreement with both AO and full results for  $E_t > 10$  GeV, while the primary-radiation contribution exhibit relevant discrepancies. As for PYTHIA, the new model, ordered in transverse momentum, is in good agreement with the resummation, leading to results similar to HERWIG. On the contrary, a visible disagreement is present between the old PYTHIA model and the resummed curves. In fact, as discussed in [9], evolution in transverse momentum leads to a better treatment of angular ordering with respect to virtuality ordering. Comparing the spectra at  $E_t = 10$  GeV, the discrepancies with respect to the full resummed result amount to  $-10\%$  for HERWIG,  $+7.5\%$  for the new PYTHIA model and  $-50\%$  for the old PYTHIA.

In Fig. 3 we instead compare HERWIG, PYTHIA and the resummation for a rapidity slice  $\Delta\eta = 3$ . As in Fig. 2, HERWIG is in reasonable agreement with the resummed computation for  $E_t > 10$  GeV and the old PYTHIA model lies quite far from the other curves throughout all

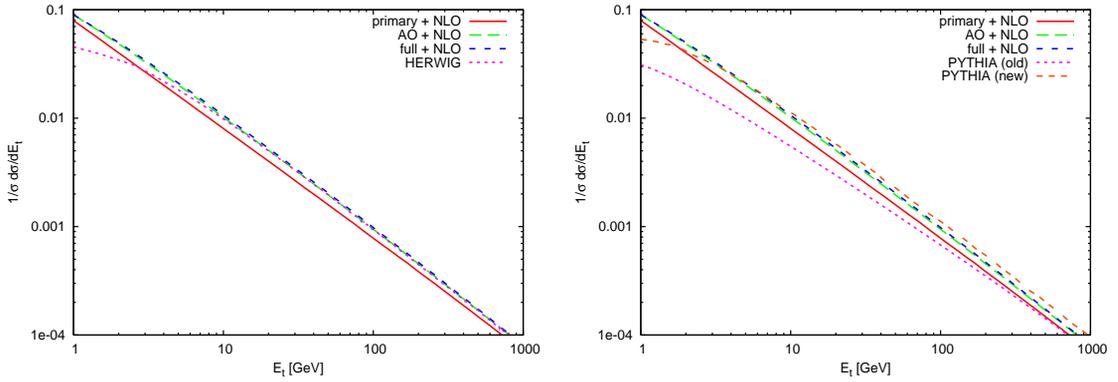


Fig. 2: Comparison of full, AO and primary resummed results with HERWIG (left) and PYTHIA (right) for  $\Delta\eta = 1$  and  $Q = 10^5$  GeV. As for PYTHIA, we show the spectra yielded by the old and new models, where parton showers are ordered in virtuality and transverse momentum, respectively.

$E_t$ -range. However, unlike the  $\Delta\eta = 1$  case, even the spectrum obtained with the new PYTHIA model exhibits a meaningful discrepancy for  $E_t > 100$  GeV, which might signal that perhaps even the new PYTHIA ordering variable is not completely adequate to describe non-global observables at large rapidity slices. A more detailed investigation of this issue is mandatory.

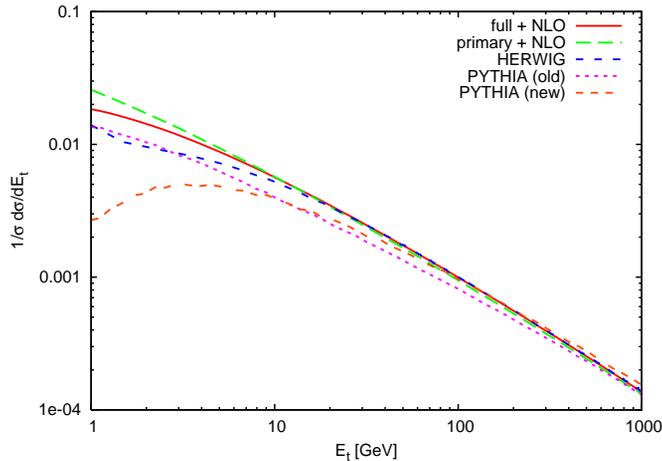


Fig. 3: Transverse-energy spectrum in a rapidity gap  $\Delta\eta = 3$ , according to the resummed calculation, HERWIG and PYTHIA (old and new models).

### 2.0.3 Conclusions

We studied non-global observables, namely the transverse-energy flow in a rapidity gap, and investigated the role played by angular ordering in the leading-logarithmic resummation. We found that the angular-ordering approximation indeed includes the bulk of the leading-logarithmic contribution, as the results are not too different with respect to the full resummed calculation.

The resummed spectra were compared with the results of the HERWIG and PYTHIA Monte Carlo generators. We found that HERWIG, whose evolution variable is equivalent to angular ordering in the soft limit, is in acceptable agreement with the resummation. As for PYTHIA, the old model, based on virtuality ordering, with an option to veto non-angular-ordered emissions, was found to be inadequate to describe non-global observables. The new model, ordered in transverse momentum and with an improved implementation of angular ordering, yields predictions qualitatively similar to HERWIG for relatively small rapidity gaps, whereas remarkable discrepancies are exhibited if the slice size is enlarged. In fact, as non-global observables are often used to tune Monte Carlo generators to data, we believe that such a discrepancy needs to be further investigated; otherwise, when fitting, *e.g.*, the old PYTHIA model to data, one would end up to include as much as 50% of perturbative leading logarithms in non-perturbative parameters, associated with hadronization or underlying event. A deeper understanding of the PYTHIA description of non-global observables, along with the application of the work here presented to hadron colliders, is in progress.

### 3 Azimuthal decorrelation between hard final state jets

*Author: Mrinal Dasgupta*

One of the most commonly measured jet observables in experimental QCD studies is the azimuthal decorrelation  $\Delta\phi$  between hard final-state jets. When compared to theory this quantity is expected to provide valuable information both on QCD parameters (strong coupling, parton distribution functions – PDFs) as well as dynamics in the near back-to-back region sensitive to multiple soft and/or collinear emissions and non-perturbative effects. To this end it has been often examined in experimental QCD studies at HERA and the Tevatron [14, 15], used for the tuning of parameters of Monte Carlo event generator models and to constrain unintegrated PDFs (uPDFs) in conjunction with HERA data [16].

In this study we aim to provide a more accurate theoretical prediction for this observable by calculating a next-to-leading logarithm (NLL) resummed result which accounts for logarithmic terms enhanced in the region where jets are back-to-back in azimuthal angle –  $\Delta\phi = \pi$ . Such a resummation has not been carried out to date, the main complication being the application of a jet algorithm to define the final state which has non-trivial implications for the standard approximations that enable NLL resummation.

To be specific one is studying here an observable that is sensitive to energy flow outside well-defined jet regions which potentially means that it and similar observables fall into the category of non-global observables [17, 18]. Since it was shown that the resummation of non-global observables is substantially more complicated than that for “global” quantities such as most event-shape variables and in any case restricted to the large  $N_c$  approximation, the most accurate theoretical predictions can be obtained only for global observables. This appears to rule out the possibility of complete NLL estimates for many interesting jet observables including the azimuthal decorrelation we study here. As far as existing predictions for jet observables are concerned, the issue of non-global logarithms was not dealt with in Ref. [19] (published prior to the discovery of non-global effects) where they would arise in threshold resummation for one of the definitions ( $M^2 = (p_1 + p_2)^2$ ) of the dijet invariant mass studied there but would be absent for the definition  $M^2 = 2p_1 \cdot p_2$ . Further we should also mention here that the non-global component has been incorrectly treated in Ref. [20] where it is mentioned that such effects will vanish with

jet radius when in fact one obtains a saturation in the small  $R$  limit.

We shall show an interplay between the potential non-global nature of the observable and the exact definition of the jet as provided by the choice of a recombination scheme. We show that in one of the resummation schemes employed in experimental studies of the azimuthal correlation the observable is in fact global and can be resummed to NLL accuracy. This may be taken as a general example of how carefully selecting the definition of the observable and the jets one may be able to render an exact NLL resummation possible, avoiding altogether the non-global issue and hence encourage future resummed studies for important regions of phase space in the context of jets.

### 3.0.4 Recombination scheme, kinematics and globalness

We wish to study the impact of two recombination schemes used to construct the angle  $\Delta\phi$  between the final-state jets in dijet production. In the first scheme the jet azimuthal angle  $\phi_j$  is given by a  $p_t$ -weighted sum over its hadronic constituents,  $\phi_j = \sum_{i \in j} p_{t,i} \phi_i / \sum_{i \in j} p_{t,i}$ , while in the second scheme one constructs the jet four-vector  $p_j = \sum_{i \in j} p_i$ , with the sum running over hadrons in the jet, and then parameterises  $p_j = p_{t,j} (\cosh \eta_j, \cos \phi_j, \sin \phi_j, \sinh \eta_j)$  to obtain the jet azimuth  $\phi_j$ . The first scheme is employed for instance by the H1 collaboration at HERA while the latter ( $E$ -scheme) is currently preferred by the Tevatron experiments.

The transverse momenta of final-state particles can be parameterised as below:<sup>1</sup>

$$\begin{aligned} \vec{p}_{t,1} &= p_{t,1}(1, 0), \\ \vec{p}_{t,2} &= p_{t,2}(\cos(\pi - \epsilon), \sin(\pi - \epsilon)), \\ &= p_{t,2}(-\cos \epsilon, \sin \epsilon), \\ \vec{k}_{t,i} &= k_{t,i}(\cos \phi_i, \sin \phi_i), \end{aligned} \tag{5}$$

where the hard final-state partons are labeled by 1 and 2 and the soft gluons by the label  $i$ . For only soft emissions the hard partons are nearly back-to-back,  $p_{t,1} = p_{t,2} = p_t$  and  $|\epsilon| \ll 1$ .

Using the above, in the scheme involving the  $p_t$ -weighted sum one obtains for  $\Delta\phi = \phi_{j1} - \phi_{j2}$ :

$$|\pi - \Delta\phi| = \left| \sum_i \frac{k_{t,i}}{p_t} (\sin \phi_i - \theta_{i1} \phi_i - \theta_{i2}(\pi - \phi_i)) \right| + \mathcal{O}(k_t^2), \tag{6}$$

where  $\theta_{ij} = 1$  if particle  $i$  is clustered to jet  $j$  and is zero otherwise. The definition above implies that the observable in question is global since it is sensitive to soft emissions in the whole phase-space, both in and outside the jets, and the dependence on soft emissions in either case is linear in  $k_t$ . This property ensures that it is possible to resum the large logarithms in the back-to-back region to next-to-leading (single) logarithmic accuracy without resorting to the large  $N_c$  approximation needed for non-global observables [17, 18].

Now turning to the  $E$ -scheme one obtains instead:

$$|\pi - \Delta\phi| = \left| \sum_{i \notin \text{jets}} \frac{k_{t,i}}{p_t} \sin \phi_i \right| + \mathcal{O}(k_t^2), \tag{7}$$

---

<sup>1</sup>Here one is looking at the projections of particle momenta in the plane perpendicular to the beam direction in hadron collisions or that perpendicular to the  $\gamma^*P$  axis in the DIS Breit or hadronic centre-of-mass (HCM) frames.

where the sum extends only over all soft particles not recombined with the hard jets. Observables sensitive to soft emissions in such delimited angular intervals are of the non-global variety [17, 18], and hence in the  $E$ -scheme definition of jets the azimuthal decorrelation is a non-global observable.

### 3.0.5 Resummed Results

Having established that the observable at hand is a global observable in the  $p_t$ -weighted recombination scheme its resummation is now straightforward. We refer the reader to Ref. [21] for the details and just quote the results below.

Taking first the case of dijets produced in DIS, the integrated cross-section ie the integral of the distribution in  $\pi - \Delta\phi$  up to some fixed value  $\Delta$  is given by an integral over ‘‘impact parameter’’  $b$

$$\Sigma_a(\Delta) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{db}{b} \sin(b\Delta) e^{-R_a(b)} f_a(x, \mu_f^2/b^2). \quad (8)$$

The index  $a$  denotes the flavour of incoming parton and the function  $R_a(b)$ , known as the radiator, embodies the soft and/or collinear single-gluon result for emission from a three hard parton system while  $f$  denotes the PDF.

For the case of hadron collisions one can write a very similar formula to the one above except that in this case one has to account for two incoming partons and hence there are two PDFs while the relevant radiator now represents soft and collinear resummation from an ensemble of four hard partons.

The result for  $R_a(\bar{b})$  for the DIS case can be expressed in terms of three pieces each with a distinct physical origin:

$$R_a(\bar{b}) = R_{\text{in}}^a(\bar{b}) + R_{\text{out}}^a(\bar{b}) - \ln S(\bar{b}, \{p\}), \quad (9)$$

with  $R_{\text{in}}^a$  and  $R_{\text{out}}^a$  being the contributions generated by emissions collinear to the incoming (excluding the set of single-logarithms already resummed in the parton densities) and outgoing legs respectively. In addition to these jet functions we have a soft function  $S(\bar{b}, \{p\})$  which resums soft emission at large angles, and which depends on the geometry of the emitting hard ensemble expressed here as a dependence on the set of hard Born momenta  $\{p\}$ .

While our results eventually include the two-loop running of the coupling which is necessary to obtain full NLL accuracy (compute the full functions  $g_1$  and  $g_2$ ), for brevity and to illustrate the main features we report our results here in a fixed coupling approximation. In this case we simply obtain:

$$R_{\text{out}}^a(\bar{b}) = (C_1^a + C_2^a) \frac{\alpha_s}{2\pi} \left( \frac{2}{3} L^2 + \frac{4}{3} L \left( -\ln 3 - 4 \ln 2 + 3 \ln \frac{Q}{p_t} \right) \right) + \frac{4}{3} \frac{\alpha_s}{2\pi} (C_1^a B_1^a + C_2^a B_2^a) L, \quad (10)$$

$$R_{\text{in}}^a = C_i^a \frac{\alpha_s}{2\pi} \left( 2L^2 + 4L \left( -\ln 2 + \ln \frac{Q}{p_t} \right) \right) + 4C_i^a \frac{\alpha_s}{2\pi} B_i^a L, \quad (11)$$

$$\ln S(\bar{b}, \{p\}) = -4L \left( 2C_F \frac{\alpha_s}{2\pi} \ln \frac{Q_{qq'}}{Q} + C_A \frac{\alpha_s}{2\pi} \ln \frac{Q_{qq} Q_{gg'}}{Q_{qq'} Q} \right), \quad (12)$$

with  $L = \ln \bar{b}$ . In the above  $C_i^a$  is the colour charge of the incoming parton in channel  $a$ , for instance  $C_i^a = C_F$  for  $a = q$ , the incoming quark channel. Likewise  $C_{1,2}^a$  are the colour charges of the partons initiating the outgoing jets 1 and 2 in channel  $a$ . The main aspect of the results for the collinear  $R_{\text{out,in}}^a$  jet functions is a leading double logarithmic behaviour, where one notes the unfamiliar coefficient  $2/3$  (different from all commonly studied event shape variables for instance) associated to the double logs on the outgoing legs, *i.e.* in the function  $R_{\text{out}}^a$ . Additionally hard collinear radiation is described by single-logarithmic terms with the coefficients  $C_\ell B_\ell$  for each leg, with the appropriate colour charge  $C_\ell$  ( $\ell = i, 1, 2$ ) and  $B_{i,1,2}$  depending on the identities (spins) of the incoming and outgoing partons such that  $B_\ell = -3/4$  for fermions and  $B_\ell = -(11C_A - 4T_R n_f)/(12C_A)$  for a gluon.

Finally we have the soft wide-angle single-logarithmic contribution  $\ln S$ , which depends on the geometry of the hard three-jet system via the dependence on dipole invariant masses  $Q_{ij} = 2(p_i \cdot p_j)$ . This structure is characteristic of soft inter-jet radiation for three-jet systems (see *e.g.* Ref. [22] for a detailed discussion). The result can be easily extended to the case of hadron collisions as shown in Ref. [21].

### 3.0.6 Results and Discussion

To provide a final resummed result for the  $\Delta\phi$  distribution one still needs to carry out the  $b$  integration in Eq. (8). The  $b$  integral is not well behaved at small and large  $b$ . At small  $b$  one is outside the jurisdiction of resummation and hence free to modify the small  $b$  behaviour with a prescription that does not affect the next-to-leading logarithms (see Ref. [21]). At large  $b$  one has to regulate the effect of the Landau pole in the running coupling and introduce non-perturbative corrections which procedure is described in Ref. [21].

We plot the resummed result for the  $\Delta\phi$  distribution in Fig. 4 along with the fixed order predictions for dijet production in DIS with  $Q^2 = 67 \text{ GeV}^2$  and  $x = 2.86 \cdot 10^{-3}$ . These values and other cuts on the jets have been taken from the H1 study to which we would eventually compare our results. As we can see the fixed order predictions diverge as expected near  $\Delta\phi = \pi$ . This divergence is cured by the resummation that goes to a fixed *non-zero* value at  $\Delta\phi = \pi$ . Of note here is the absence of a Sudakov peak since the Sudakov mechanism does not dominate the  $b$  integral at very small  $\Delta = |\pi - \Delta\phi|$ . The dominant mechanism to obtain back-to-back jets is thus a one-dimensional cancellation between emissions rather than a suppression of the  $k_t$  of each individual emission, leading to a washout of the Sudakov peak.

In order to obtain complete predictions which can be compared to data two further developments need to be made: matching to fixed-order NLO predictions and inclusion of non-perturbative effects. These issues will be addressed in forthcoming work.

## 4 Matching of NLLA to NNLO calculation for event shapes in $e^+e^-$

*Author: Gionata Luisoni*

Event shape distributions in  $e^+e^-$  annihilation processes are very popular hadronic observables. Their popularity is mainly due to the fact that they are well suited both for experimental measurement and for theoretical calculations because many of them are infrared and collinear safe.

The deviation from simple two-jet configurations, which are a limiting case in event shapes, is proportional to the strong coupling constant  $\alpha_s$ , so that by comparing the measured

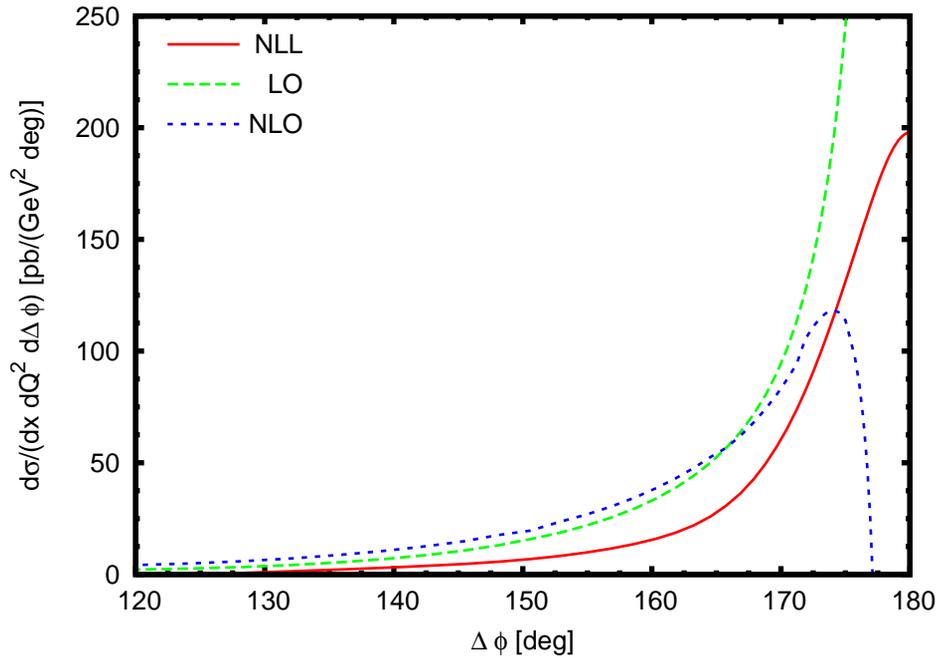


Fig. 4: The resummed  $\Delta\phi$  distribution for dijets in DIS. Also shown for comparison are the leading order (LO) and next-to-leading order (NLO) predictions from NLOJET++.

$\bar{\alpha}_s \mathcal{A}(y)$	$\bar{\alpha}_s L$	$\bar{\alpha}_s L^2$				
$\bar{\alpha}_s^2 \mathcal{B}(y, x_\mu)$	$\bar{\alpha}_s^2 L$	$\bar{\alpha}_s^2 L^2$	$\bar{\alpha}_s^2 L^3$	$\bar{\alpha}_s^2 L^4$		
$\bar{\alpha}_s^3 \mathcal{C}(y, x_\mu)$	$\bar{\alpha}_s^3 L$	$\bar{\alpha}_s^3 L^2$	$\bar{\alpha}_s^3 L^3$	$\bar{\alpha}_s^3 L^4$	$\bar{\alpha}_s^3 L^5$	$\bar{\alpha}_s^3 L^6$

Table 1: Powers of the logarithms present at different orders in perturbation theory. The colour highlights the different orders in resummation: LL (red) and NLL (blue). The terms in green are contained in the LL and NLL contributions and exponentiate trivially with them.

event shape distribution with the theoretical prediction, one can determine  $\alpha_s$  [23]. Below we will concentrate on this, using the newly available NNLO [24] and NLLA+NNLO results. At LEP, a standard set of event shapes was studied in great detail: thrust  $T$  (which is substituted here by  $\tau = 1 - T$ ), heavy jet mass  $\rho$ , wide and total jet broadening  $B_W$  and  $B_T$ ,  $C$ -parameter and two-to-three-jet transition parameter in the Durham algorithm  $Y_3$ . The definitions of these variables, which we denote collectively as  $y$  in the following, are summarized in [25]. The two-jet limit of each variable is  $y \rightarrow 0$ .

The theoretical state-of-the-art description of event shape distributions was based until very recently on the matching of the NLLA [26] onto the NLO [27–30] calculation. The newly available results of the NNLO corrections for the standard set of event shapes [24] introduced above, permits now to match them with resummed calculations, obtaining theoretical distributions at NLLA+NNLO.

At NNLO the integrated cross section

$$R(y, Q, \mu) \equiv \frac{1}{\sigma_{\text{had}}} \int_0^y \frac{d\sigma(x, Q, \mu)}{dx} dx,$$

has the following fixed-order expansion:

$$R(y, Q, \mu) = 1 + \bar{\alpha}_s(\mu) \mathcal{A}(y) + \bar{\alpha}_s^2(\mu) \mathcal{B}(y, x_\mu) + \bar{\alpha}_s^3(\mu) \mathcal{C}(y, x_\mu).$$

where  $\bar{\alpha}_s = \alpha_s/(2\pi)$  and  $x_\mu = \mu/Q$ . Approaching the two-jet region event shapes display large infrared logarithms which spoil the convergence of the series expansion. The main contribution in this case comes from the highest power of the logarithms which have to be resummed to all orders. For suitable observables resummation leads to exponentiation. At NLLA the resummed expression is given by

$$R(y, Q, \mu) = (1 + C_1 \bar{\alpha}_s) e^{(L g_1(\alpha_s L) + g_2(\alpha_s L))},$$

where the function  $g_1(\alpha_s L)$  contains all leading-logarithms (LL),  $g_2(\alpha_s L)$  all next-to-leading-logarithms (NLL) and  $\mu = Q$  is used. Terms beyond NLL have been consistently omitted. The resummation functions  $g_1(\alpha_s L)$  and  $g_2(\alpha_s L)$  can be expanded as power series in  $\bar{\alpha}_s L$

$$\begin{aligned} L g_1(\alpha_s L) &= G_{12} L^2 \bar{\alpha}_s + G_{23} L^3 \bar{\alpha}_s^2 + G_{34} L^4 \bar{\alpha}_s^3 + \dots \text{ (LL)}, \\ g_2(\alpha_s L) &= G_{11} L \bar{\alpha}_s + G_{22} L^2 \bar{\alpha}_s^2 + G_{33} L^3 \bar{\alpha}_s^3 + \dots \text{ (NLL)}. \end{aligned} \quad (13)$$

Table 1 shows the logarithmic terms present up to the third order in perturbation theory. From the expansion (13) of the exponentiated resummation functions it follows immediately, that at

the fixed-order level, the LL are terms of the form  $\alpha_s^n L^{n+1}$ , the NLL terms go like  $\alpha_s^n L^n$ , and so on.

Closed analytic forms for functions  $g_1(\alpha_s L)$ ,  $g_2(\alpha_s L)$  are available for  $\tau$  and  $\rho$  [31],  $B_W$  and  $B_T$  [32, 33],  $C$  [34, 35] and  $Y_3$  [36], and are collected in the appendix of [37]. For  $\tau$  the  $g_3(\alpha_s L)$  function is also known [38].

To obtain a reliable description of the event shape distributions over a wide range in  $y$ , it is mandatory to combine fixed-order and resummed predictions. To avoid the double counting of terms common to both, the two predictions have to be matched to each other. A number of different matching procedures have been proposed in the literature, see for example [25] for a review. We computed the matching in the so-called  $\ln R$ -matching [26] since in this particular scheme, all matching coefficients can be extracted analytically from the resummed calculation, while most other schemes require the numerical extraction of some of the matching coefficients from the distributions at fixed order. The  $\ln R$ -matching at NLO is described in detail in [26]. In the  $\ln R$ -matching scheme, the NLLA+NNLO expression is

$$\begin{aligned} \ln(R(y, \alpha_S)) &= L g_1(\alpha_S L) + g_2(\alpha_S L) + \bar{\alpha}_S (\mathcal{A}(y) - G_{11}L - G_{12}L^2) \\ &\quad + \bar{\alpha}_S^2 \left( \mathcal{B}(y) - \frac{1}{2} \mathcal{A}^2(y) - G_{22}L^2 - G_{23}L^3 \right) \\ &\quad + \bar{\alpha}_S^3 \left( \mathcal{C}(y) - \mathcal{A}(y) \mathcal{B}(y) + \frac{1}{3} \mathcal{A}^3(y) - G_{33}L^3 - G_{34}L^4 \right). \end{aligned} \quad (14)$$

The matching coefficients appearing in this expression can be obtained from (13) and are listed in [37]. To ensure the vanishing of the matched expression at the kinematical boundary  $y_{\max}$  a further shift of the logarithm is made [25].

The full renormalisation scale dependence of (14) is given by replacing the coupling constant, the fixed-order coefficients, the resummation functions and the matching coefficients as follows:

$$\begin{aligned} \alpha_s &\rightarrow \alpha_s(\mu), \\ \mathcal{B}(y) &\rightarrow \mathcal{B}(y, \mu) = 2\beta_0 \ln x_\mu \mathcal{A}(y) + \mathcal{B}(y), \\ \mathcal{C}(y) &\rightarrow \mathcal{C}(y, \mu) = (2\beta_0 \ln x_\mu)^2 \mathcal{A}(y) + 2 \ln x_\mu [2\beta_0 \mathcal{B}(y) + 2\beta_1 \mathcal{A}(y)] + \mathcal{C}(y), \\ g_2(\alpha_S L) &\rightarrow g_2(\alpha_S L, \mu^2) = g_2(\alpha_S L) + \frac{\beta_0}{\pi} (\alpha_S L)^2 g_1'(\alpha_S L) \ln x_\mu, \\ G_{22} &\rightarrow G_{22}(\mu) = G_{22} + 2\beta_0 G_{12} \ln x_\mu, \\ G_{33} &\rightarrow G_{33}(\mu) = G_{33} + 4\beta_0 G_{23} \ln x_\mu. \end{aligned}$$

In the above,  $g_1'$  denotes the derivative of  $g_1$  with respect to its argument. The LO coefficient  $\mathcal{A}$  and the LL resummation function  $g_1$ , as well as the matching coefficients  $G_{i i+1}$  remain independent of  $\mu$ .

In the two upper plots of Fig. 5 we compare the matched NLLA+NNLO predictions for the heavy jet mass with the fixed-order NNLO predictions, and the matched NLLA+NLO with fixed-order NLO. All distributions were weighted by the respective shape variables. We use  $Q = M_Z$  and fix  $x_\mu = 1$ , the strong coupling constant is taken as  $\alpha_s(M_Z) = 0.1189$ . To quantify the renormalisation scale uncertainty, we have varied  $1/2 < x_\mu < 2$ , resulting in the error

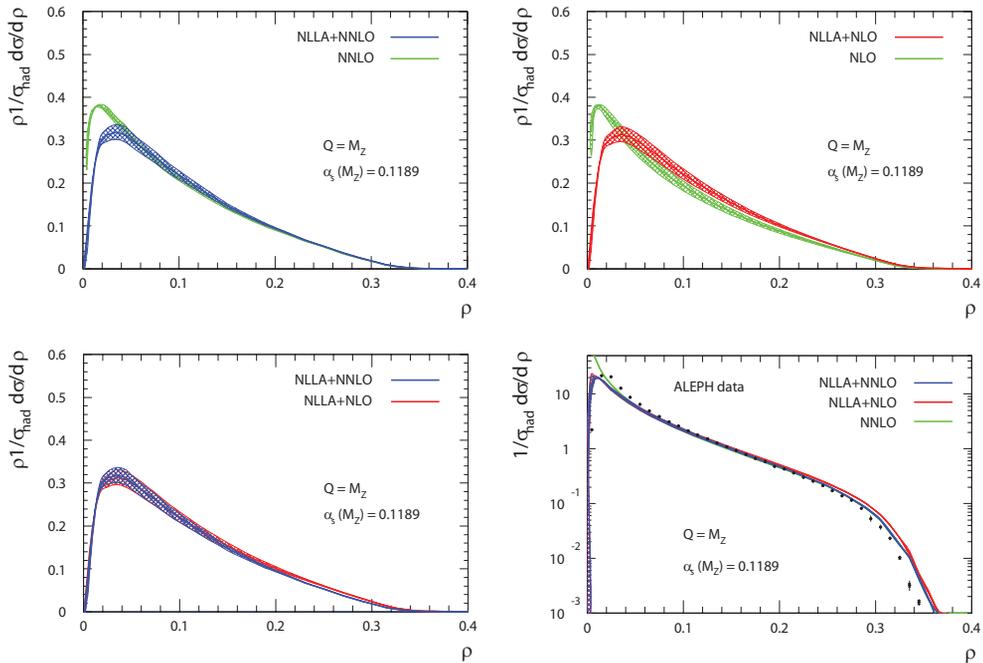


Fig. 5: Matched distributions of heavy jet mass  $\rho$ .

band on these figures. The effects visible for the heavy jet mass are common to the whole set of observables which were analyzed. The most striking observation is that the difference between NLLA+NNLO and NNLO is largely restricted to the two-jet region, while NLLA+NLO and NLO differ in normalisation throughout the full kinematical range. This behaviour may serve as a first indication for the numerical smallness of corrections beyond NNLO in the three-jet region. In the approach to the two-jet region, the NLLA+NLO and NLLA+NNLO predictions agree by construction, since the matching suppresses any fixed-order terms. On the plot in the lower left corner we observe that the difference between NLLA+NNLO and NLLA+NLO is only moderate in the three-jet region. The renormalisation scale uncertainty in the three-jet region is reduced by 20-40% between NLLA+NLO and NLLA+NNLO. Finally the lower-right plot shows the parton-level fixed NNLO and the matched NLLA+NLO and NLLA+NNLO predictions are compared to hadron-level data taken by the ALEPH experiment. The description of the hadron-level data improves between parton-level NLLA+NLO and parton-level NLLA+NNLO, especially in the three-jet region. The behavior in the two-jet region is described better by the resummed predictions than by the fixed-order NNLO, although the agreement is far from perfect. This discrepancy can in part be attributed to hadronisation corrections, which become large in the approach to the two-jet limit. A very recent study of logarithmic corrections beyond NLLA for the thrust distribution [38] also shows that subleading logarithms in the two-jet region can account for about half of this discrepancy.

With the new NNLO and NLLA+NNLO results a new extraction of  $\alpha_s$  can be performed. For this we used public ALEPH data at center-of-mass energies between 91 and 209 GeV [39]. The data are corrected to hadron level using Monte Carlo (MC) corrections and accounting for

initial- and final-state-radiation (ISR/FSR) as well as background. They are fitted by NNLO respectively NLLA+NNLO predictions, including NLO quark mass corrections, folded to hadron level by means of MC generators. Finally, after estimating the missing higher orders using the uncertainty band method [25], the fits of 8 data sets and 6 different variables are combined together [23].

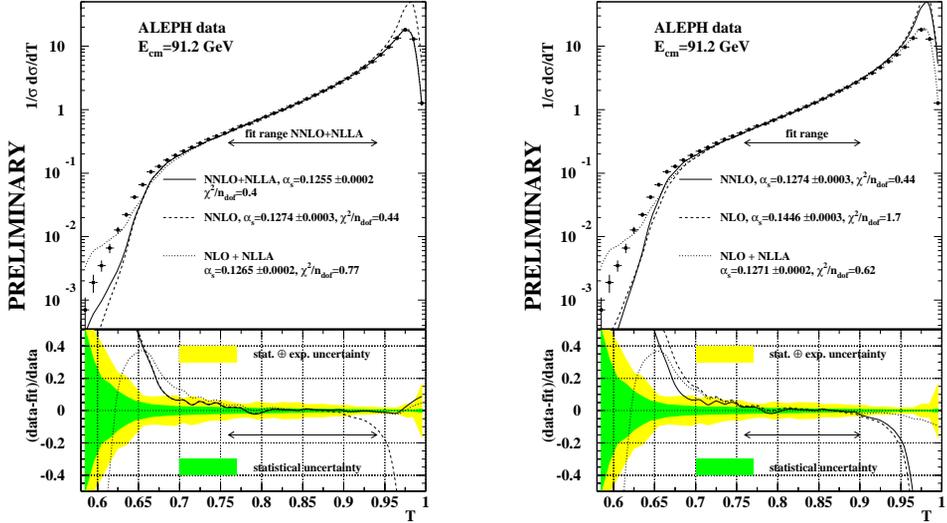


Fig. 6: Fit to ALEPH data for thrust.

The part of the distribution chosen for the fit (Fig. 6) is the one where the hadronizations and detector corrections are smaller than 25%. In the case of the NNLO distributions, the range was further reduced in the 2-jet region because of the divergence of the theoretical predictions. Only the statistical uncertainties are included in the  $\chi^2$ .

At NNLO we see a clear improvement with respect to the old NLO results. The fit is of a good quality although it still includes large statistical uncertainties of the  $\mathcal{C}$  coefficient and in the 2-jet region the NLLA+NLO predictions still yields a better result. The improvement between NNLO and NLLA+NNLO is visible especially in the 2-jet region. The fit range is also more extended in this direction. For the resulting  $\alpha_s$  we observe that using fixed-order predictions leads basically to higher values, and that in both fixed-order and matched predictions there is a tendency for  $\alpha_s$  to decrease passing from NLO to NNLO. Finally computing the weighted average for  $\alpha_s$  from the 6 variables we obtain [23]:

$$\bar{\alpha}_s(M_Z) = 0.1240 \pm 0.0008 (\text{stat}) \pm 0.0010 (\text{exp}) \pm 0.0011 (\text{had}) \pm 0.0029 (\text{theo}).$$

From Fig. 7 it is clearly visible that the results for the different variables are coherent and the scattering is much reduced. The improvement with respect to the NLO result is also remarkable.

The combined results for the NLLA+NNLO fits are still work-in-progress, but it can be anticipated that the improvement coming from the inclusion of resummed calculation will be less

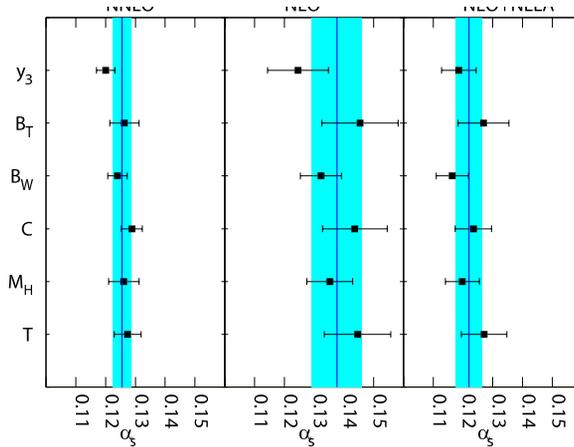


Fig. 7: Combination of  $\alpha_s$  fits at NLO, NLLA+NLO and NNLO.

dramatic than the one obtained at NLO level. The reason for this is that the compensation for the two-loop running of the coupling constant is present only in the NNLO coefficient and not in the resummed part.

These results show that there is space for further improvements, which could be obtained by resumming subleading logarithms similarly to what was done recently for thrust [38]. Improvements are also expected from the addition of electroweak corrections. Finally, a further step forward in the comparison of theoretical predictions with experimental data could be done by using modern MC tools based on NLO calculations matched with parton showers for the computation of the hadronization corrections.

## 5 Precision resummed QEDxQCD theory for LHC physics: status and update

*Authors: B.F.L. Ward, S. Joseph, Swapan Majhi, S.A. Yost*

With the advent of the LHC, we enter the era of precision QCD, by which we mean predictions for QCD processes at the total theoretical precision tag of 1% or better. The attendant requirement for this theoretical precision is control of the  $\mathcal{O}(\alpha_s^2 L^{n_1}, \alpha_s \alpha L^{n_2}, \alpha^2 L^{n_3})$ ,  $n_1 = 0, 1, 2$ ,  $n_2 = 1, 2$ ,  $n_3 = 2$  corrections in the presence of realistic parton showers, on an event-by-event basis – here,  $L$  is a generic big logarithm. This is the objective of our approach to precision QCD theory, which for example will be needed for the expected 2% experimental precision [40–42] at the LHC for processes such as  $pp \rightarrow V + m(\gamma) + n(G) + X \rightarrow \bar{\ell}\ell' + m'(\gamma) + n(G) + X$ ,  $V = W^\pm, Z$ , and  $\ell = e, \mu$ ,  $\ell' = \nu_e, \nu_\mu(e, \mu)$  for  $V = W^+(Z)$  respectively, and  $\ell = \nu_e, \nu_\mu$ ,  $\ell' = e, \mu$  respectively for  $V = W^-$ . Here, we present the elements of our approach and its recent applications in Monte Carlo (MC) event generator studies, which are still preliminary.

At such a precision as we have as our goal, issues such as the role of QED are an integral part of the discussion and we deal with this by the simultaneous resummation of QED and QCD large infrared (IR) effects,  $QED \otimes QCD$  resummation [43–49] in the presence of parton showers, to be realized on an event-by-event basis by MC methods. This is reviewed in the next section. Let us note already that in Refs. [50–55] it has been shown that QED evolution enters

at the  $\sim 0.3\%$  level for parton distributions and that in Refs. [56, 57] it has been shown that EW (large Sudakov logs, etc.) effects at LHC energies, as W's and Z's are almost massless on the TeV scale, can enter at the several percent level – such corrections must be treated systematically before any claim of 1% precision can be taken seriously. We are presenting a framework in which this can be done. The new amplitude-based resummation algebra then leads to a new scheme for calculating hard hadron-hadron scattering processes, IR-improved DGLAP-CS theory [58] for parton distributions, kernels, reduced cross sections with the appropriate shower/ME matching. This is summarized in Sec. 1.4.3. In this latter section, with an eye toward technical precision cross checks plus possible physical effects of heavy quark masses, we also deal with the issue of quark masses as collinear regulators [59–63] as an alternative [64] to the usual practice of setting all initial state quark masses to zero in calculating initial state radiation (ISR) effects in higher order QCD corrections. We also discuss in Sec. 1.4.3 the relationship between our resummation algebra and that of Refs. [65–69], as again such comparisons will be necessary in assessing the ultimate theoretical precision tag. In Sec. 1.4.4, we illustrate recent results we have obtained for the effects of our new approach on the parton showers as they are generated with the HERWIG6.5 MC [70]. Extensions of such studies to PYTHIA [71] and MC@NLO [72, 73] are in progress. Section 1.4.5 contains summary remarks.

As a point of reference, in Ref. [74] it has been argued that the current state-of-the-art theoretical precision tag on single Z production at the LHC is  $(4.1 \pm 0.3)\% = (1.51 \pm 0.75)\%(QCD) \oplus 3.79(PDF) \oplus 0.38 \pm 0.26(EW)\%$ , where the results of Refs. [72, 73, 75–86] have been used in this precision tag determination.<sup>2</sup>

### 5.0.7 QED $\otimes$ QCD Resummation

In Refs. [43–49], we have extended the YFS theory to the simultaneous exponentiation of the large IR terms in QCD and the exact IR divergent terms in QED, so that for the prototypical subprocesses  $\bar{Q}'Q \rightarrow \bar{Q}'''Q'' + m(G) + n(\gamma)$  we arrive at the new result

$$\begin{aligned}
 d\hat{\sigma}_{\text{exp}} &= e^{\text{SUM}_{\text{IR}}(\text{QCED})} \\
 &\sum_{m,n=0}^{\infty} \frac{1}{m!n!} \int \prod_{j_1=1}^m \frac{d^3 k_{j_1}}{k_{j_1}} \prod_{j_2=1}^n \frac{d^3 k'_{j_2}}{k'_{j_2}} \int \frac{d^4 y}{(2\pi)^4} \\
 &e^{iy \cdot (p_1 + q_1 - p_2 - q_2 - \sum k_{j_1} - \sum k'_{j_2}) + D_{\text{QCED}}} \\
 &\tilde{\beta}_{m,n}(k_1, \dots, k_m; k'_1, \dots, k'_n) \frac{d^3 p_2}{p_2^0} \frac{d^3 q_2}{q_2^0},
 \end{aligned} \tag{15}$$

where the new YFS [88–98] residuals, defined in Refs. [43–49],  $\tilde{\beta}_{m,n}(k_1, \dots, k_m; k'_1, \dots, k'_n)$ , with  $m$  hard gluons and  $n$  hard photons, represent the successive application of the YFS expansion first for QCD and subsequently for QED. The functions  $\text{SUM}_{\text{IR}}(\text{QCED})$ ,  $D_{\text{QCED}}$  are determined from their analogues  $\text{SUM}_{\text{IR}}(\text{QCD})$ ,  $D_{\text{QCD}}$  in Refs. [99–104] via the substitutions

$$\begin{aligned}
 B_{\text{QCD}}^{\text{nls}} &\rightarrow B_{\text{QCD}}^{\text{nls}} + B_{\text{QED}}^{\text{nls}} \equiv B_{\text{QCED}}^{\text{nls}}, \\
 \tilde{B}_{\text{QCD}}^{\text{nls}} &\rightarrow \tilde{B}_{\text{QCD}}^{\text{nls}} + \tilde{B}_{\text{QED}}^{\text{nls}} \equiv \tilde{B}_{\text{QCED}}^{\text{nls}}, \\
 \tilde{S}_{\text{QCD}}^{\text{nls}} &\rightarrow \tilde{S}_{\text{QCD}}^{\text{nls}} + \tilde{S}_{\text{QED}}^{\text{nls}} \equiv \tilde{S}_{\text{QCED}}^{\text{nls}}
 \end{aligned} \tag{16}$$

<sup>2</sup>Recently, the analogous estimate for single W production has been given in Ref. [87] – it is  $\sim 5.7\%$ .

everywhere in expressions for the latter functions given in Refs. [99–104] – see Refs. [43–49] for the details of this substitution. It can be readily established [43–49] that the QCD dominant corrections happen an order of magnitude earlier in time compared to those of QED so that the leading term  $\tilde{\beta}_{0,0}$  already gives us a good estimate of the size of the effects we study.

Important in any total theoretical prediction is knowledge of possible systematic issues associated with one’s methods. This entails the relationship between different approaches to the same classes of corrections and moves us to the relationship between our approach to QCD resummation and the more familiar approach in Refs. [65–67]. It has been shown in Ref. [105] that the latter approach is entirely equivalent to the approach in Refs. [68, 69]. Establishing the relationship between our approach and that in Refs. [65–67] will then suffice to relate all three approaches.

In Ref. [106] the more familiar resummation for soft gluons in Refs. [65–67] is applied to a general  $2 \rightarrow n$  parton process [f] at hard scale  $Q$ ,  $f_1(p_1, r_1) + f_2(p_2, r_2) \rightarrow f_3(p_3, r_3) + f_4(p_4, r_4) + \dots + f_{n+2}(p_{n+2}, r_{n+2})$ , where the  $p_i, r_i$  label 4-momenta and color indices respectively, with all parton masses set to zero to get

$$\begin{aligned} \mathcal{M}_{\{r_i\}}^{[f]} &= \sum_L^C \mathcal{M}_L^{[f]}(c_L)_{\{r_i\}} \\ &= J^{[f]} \sum_L^C S_{LI} H_I^{[f]}(c_L)_{\{r_i\}}, \end{aligned} \quad (17)$$

where repeated indices are summed,  $J^{[f]}$  is the jet function,  $S_{LI}$  is the soft function which describes the exchange of soft gluons between the external lines, and  $H_I^{[f]}$  is the hard coefficient function. The attendant IR and collinear poles are calculated to 2-loop order. To make contact with our approach, identify in  $\bar{Q}'Q \rightarrow \bar{Q}'''Q'' + m(G)$  in (15)  $f_1 = Q, \bar{Q}' = \bar{Q}', f_2 = \bar{Q}', f_3 = Q'', f_4 = \bar{Q}''', \{f_5, \dots, f_{n+2}\} = \{G_1, \dots, G_m\}$  so that  $n = m+2$  here. Observe the following:

- By its definition in Eq.(2.23) of Ref. [106], the anomalous dimension of the matrix  $S_{LI}$  does not contain any of the diagonal effects described by our infrared functions  $\Sigma_{IR}(QCD)$  and  $D_{QCD}$ .
- By its definition in Eqs.(2.5) and (2.7) of Ref. [106], the jet function  $J^{[f]}$  contains the exponential of the virtual infrared function  $\alpha_s \mathfrak{R}B_{QCD}$ , so that we have to take care that we do not double count when we use (17) in (15) and the equations that lead thereto.

It follows that, referring to our analysis in Ref. [107], we identify  $\bar{\rho}^{(m)}$  in Eq.(73) in this latter reference in our theory as

$$\begin{aligned} \bar{\rho}^{(m)}(p_1, q_1, p_2, q_2, k_1, \dots, k_m) &= \overline{\sum_{colors, spin} |\mathcal{M}_{\{r_i\}}^{\prime [f]}|^2} \\ &\equiv \sum_{spins, \{r_i\}, \{r'_i\}} \mathfrak{h}_{\{r_i\}, \{r'_i\}}^{cs} |\bar{J}^{[f]}|^2 \sum_{L=1}^C \sum_{L'=1}^C S_{LI}^{[f]} H_I^{[f]}(c_L)_{\{r_i\}} \left( S_{L'I'}^{[f]} H_{I'}^{[f]}(c_{L'})_{\{r'_i\}} \right)^\dagger, \end{aligned} \quad (18)$$

where here we defined  $\bar{J}^{[f]} = e^{-\alpha_s \mathfrak{R}B_{QCD}} J^{[f]}$ , and we introduced the color-spin density matrix for the initial state,  $\mathfrak{h}^{cs}$ . Here, we recall (see Refs. [58, 107], for example) that in our theory, we

have

$$d\hat{\sigma}^n = \frac{e^{2\alpha_s \text{Re}B_{QCD}}}{n!} \int \prod_{m=1}^n \frac{d^3 k_m}{(k_m^2 + \lambda^2)^{1/2}} \delta(p_1 + q_1 - p_2 - q_2 - \sum_{i=1}^n k_i) \bar{\rho}^{(n)}(p_1, q_1, p_2, q_2, k_1, \dots, k_n) \frac{d^3 p_2 d^3 q_2}{p_2^0 q_2^0}, \quad (19)$$

for  $n$ -gluon emission. It follows that we can repeat thus our usual steps (see Refs. [58, 107]) to get the QCD corrections in our formula (15), without any double counting of effects. This use of the results in Ref. [106] is in progress.

### 5.0.8 IR-Improved DGLAP-CS Theory: Applications

In Refs. [58, 107] it has been shown that application of the result (15) to all aspects of the standard formula for hard hadron-hadron scattering processes,

$$\sigma = \sum_{i,j} \int dx_1 dx_2 F_i(x_1) F_j(x_2) \hat{\sigma}(x_1 x_2 s) \quad (20)$$

where the  $\{F_i(x)\}$  and  $\hat{\sigma}$  denote the parton densities and reduced cross section, respectively, leads one to its application to the DGLAP-CS theory itself for the kernels which govern the evolution of the parton densities in addition to the implied application to the respective hard scattering reduced cross section. The result is a new set of IR-improved kernels [58],

$$P_{qq}(z) = C_F F_{YFS}(\gamma_q) e^{\frac{1}{2}\delta_q} \left[ \frac{1+z^2}{1-z} (1-z)^{\gamma_q} - f_q(\gamma_q) \delta(1-z) \right], \quad (21)$$

$$P_{Gq}(z) = C_F F_{YFS}(\gamma_q) e^{\frac{1}{2}\delta_q} \frac{1+(1-z)^2}{z} z^{\gamma_q}, \quad (22)$$

$$P_{GG}(z) = 2C_G F_{YFS}(\gamma_G) e^{\frac{1}{2}\delta_G} \left\{ \frac{1-z}{z} z^{\gamma_G} + \frac{z}{1-z} (1-z)^{\gamma_G} + \frac{1}{2} (z^{1+\gamma_G} (1-z) + z(1-z)^{1+\gamma_G}) - f_G(\gamma_G) \delta(1-z) \right\}, \quad (23)$$

$$P_{qG}(z) = F_{YFS}(\gamma_G) e^{\frac{1}{2}\delta_G} \frac{1}{2} \{ z^2 (1-z)^{\gamma_G} + (1-z)^2 z^{\gamma_G} \}. \quad (24)$$

in the standard notation, where

$$\gamma_q = C_F \frac{\alpha_s}{\pi} t = \frac{4C_F}{\beta_0} \quad (25)$$

$$\delta_q = \frac{\gamma_q}{2} + \frac{\alpha_s C_F}{\pi} \left( \frac{\pi^2}{3} - \frac{1}{2} \right) \quad (26)$$

$$\gamma_G = C_G \frac{\alpha_s}{\pi} t = \frac{4C_G}{\beta_0} \quad (27)$$

$$\delta_G = \frac{\gamma_G}{2} + \frac{\alpha_s C_G}{\pi} \left( \frac{\pi^2}{3} - \frac{1}{2} \right) \quad (28)$$

and

$$F_{YFS}(\gamma_q) = \frac{e^{-C_E \gamma_q}}{\Gamma(1 + \gamma_q)}, \quad (29)$$

so that

$$f_q(\gamma_q) = \frac{2}{\gamma_q} - \frac{2}{\gamma_q + 1} + \frac{1}{\gamma_q + 2} \quad (30)$$

$$f_G(\gamma_G) = \frac{n_f}{C_G} \frac{1}{(1 + \gamma_G)(2 + \gamma_G)(3 + \gamma_G)} + \frac{2}{\gamma_G(1 + \gamma_G)(2 + \gamma_G)} \quad (31)$$

$$+ \frac{1}{(1 + \gamma_G)(2 + \gamma_G)} + \frac{1}{2(3 + \gamma_G)(4 + \gamma_G)} \quad (32)$$

$$+ \frac{1}{(2 + \gamma_G)(3 + \gamma_G)(4 + \gamma_G)}. \quad (33)$$

Here,  $C_E = 0.5772\dots$  is Euler's constant and  $\Gamma(w)$  is the Euler Gamma function. We see that the kernels are integrable at the IR end-points and this admits a more friendly MC implementation, which is in progress.

Some observations are in order. First, we note that the connection of (24) with the higher-order kernel results in Refs. [108–117] is immediate and has been shown in Refs. [58, 107]. Second, there is no contradiction with the standard Wilson expansion, as the terms we resum are not in that expansion by its usual definition. Third, we do not change the predicted cross section: we have a new scheme such that the cross section in (20) becomes

$$\sigma = \sum_{i,j} \int dx_1 dx_2 F'_i(x_1) F'_j(x_2) \hat{\sigma}'(x_1 x_2 s) \quad (34)$$

order by order in perturbation theory, where  $\{P^{exp}\}$  factorize  $\hat{\sigma}_{\text{unfactorized}}$  to yield  $\hat{\sigma}'$  and its attendant parton densities  $\{F'_i\}$ . Fourth, when one solves for the effects of the exponentiation in (24) on the actual evolution of the parton densities from the typical reference scale of  $Q_0 \sim 2\text{GeV}$  to  $Q = 100\text{ GeV}$  one finds [58, 107] shifts of  $\sim 5\%$  for the NS  $n=2$  moment for example, which is thus of some phenomenological interest— see for example Ref. [118]. Finally, we note that we have used [43–49] the result (15) for single Z production with leptonic decay at the LHC (and at FNAL) to focus on the ISR alone, for definiteness and we find agreement with the literature in Refs. [119–123] for exact  $\mathcal{O}(\alpha)$  results and Refs. [124–126] for exact  $\mathcal{O}(\alpha_s^2)$  results, with a threshold QED effect of 0.3%, similar to that found for the parton evolution itself from QED in Refs. [50–55]. Evidently, any 1% precision tag must account for all such effects.

### 5.0.9 Shower/ME Matching

In using (15) in (34) for  $\hat{\sigma}'(x_i x_j)$ , we intend to combine our exact extended YFS calculus with HERWIG [70] and PYTHIA [71] as follows: they generate a parton shower starting from  $(x_1, x_2)$  at the factorization scale  $\mu$  after this point is provided by the  $\{F'_i\}$  and we may use [43–49] either a  $p_T$ -matching scheme or a shower-subtracted residual scheme where the respective new residuals  $\{\hat{\beta}_{n,m}(k_1, \dots, k_n; k'_1, \dots, k'_m)\}$  are obtained by expanding the shower formula and the result in (15) on product and requiring the agreement with exact results to the specified order.<sup>3</sup>

<sup>3</sup>See Ref. [127, 128] for a realization of the shower subtracted residual scheme in the context of QED parton showers.

This combination of theoretical constructs can be systematically improved with exact results order-by-order in  $\alpha_s, \alpha$ , with exact phase space.<sup>4</sup> The recently developed new parton evolution algorithms in Refs. [129, 130] may also be used here.

The issue of the non-zero quark masses in the ISR is present when one wants 1% precision, as we know that the parton densities for the heavy quarks are all different and the generic size of mass corrections for bremsstrahlung is  $\alpha_s/\pi$  for cross sections [131], so that one would like to know whether regularizing a zero-mass ISR radiation result with dimensional methods, carrying through the factorization procedure gives the same result as doing the same calculation with the physical, non-zero mass of the quark and again carrying through the factorization procedure to the accuracy  $\alpha_s^2/\pi^2$ , for example. Until the analysis in Ref. [64], this cross check was not possible because in Refs. [59–62] it was shown that there is a lack of Bloch-Nordsieck cancellation in the ISR at  $\mathcal{O}(\alpha_s^2)$  unless the radiating quarks are massless. The QCD resummation algebra, as used in (15), allows us to obviate [64] this theorem, so that now such cross checks are possible and they are in progress.

#### 5.0.10 Sample MC data: IR-Improved Kernels in HERWIG6.5

We have preliminary results on IR-improved showers in HERWIG6.5: we compare the  $z$  - distributions and the  $p_T$  of the IR-improved and usual DGLAP-CS showers in the Figs. 8-10. As we would expect, the IR-improved shower re-populates the soft region in both variables. The details of the implementation procedure and the respective new version of HERWIG6.5, HERWIG6.5-YFS, will appear elsewhere [132]. The analogous implementations in PYTHIA and MC@NLO are in progress, as are comparisons with IR-safe observables.

#### 5.0.11 Conclusions

The theory of Refs. [88, 89] extends to the joint resummation of QED and QCD with proper shower/ME matching built-in. For the simultaneous QED $\otimes$ QCD resummed theory, full MC event generator realization is open: a firm basis for the complete  $\mathcal{O}(\alpha_s^2, \alpha\alpha_s, \alpha^2)$  MC results needed for precision LHC physics has been demonstrated and all the latter are in progress – see Refs. [133–137] for new results on  $\epsilon$  expansions for the higher-order Feynman integrals needed to isolate the residuals in our approach, for example. This allows cross check between residuals isolated with the quark masses as regulators, something now allowed by the result in Ref. [64], and those isolated in dimensional regularization for the massless quark limit. Such cross checks are relevant for precision QCD theory. The first MC data have been shown with IR-improved showers in HERWIG6.5. The spectra are softer as expected. We look forward to the detailed comparison with IR-safe observables as generated with IR-improved and with the usual showers – this will appear elsewhere. [132]. Already, semi-analytical results at the  $\tilde{\beta}_{0,0}^{0,0}$  are consistent with the literature on single Z production, while a cross check for the analogous W production is near. As the QED is at 0.3% at threshold, it is needed for 1% precision.

## References

- [1] G. Zanderighi et al, *Higher-order calculations*. These proceedings.

---

<sup>4</sup>The current state of the art for such shower/ME matching is given in Refs. [72, 73], which realizes exactness at  $\mathcal{O}(\alpha_s)$ .

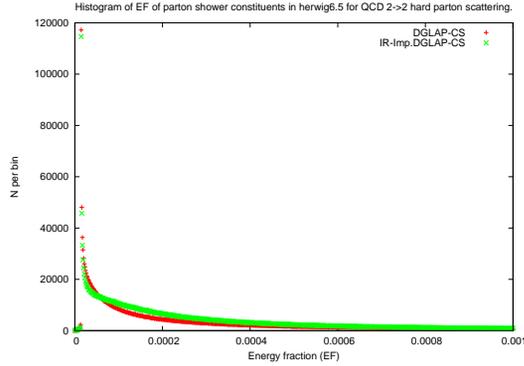


Fig. 8: The  $z$ -distribution shower comparison in HERWIG6.5 – preliminary results.

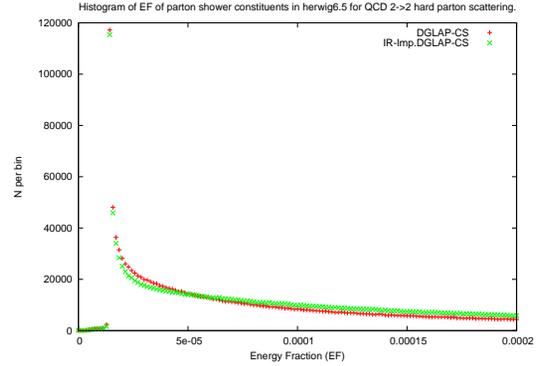


Fig. 9: The  $z$ -distribution shower comparison in HERWIG6.5 at small  $z$  – preliminary results.

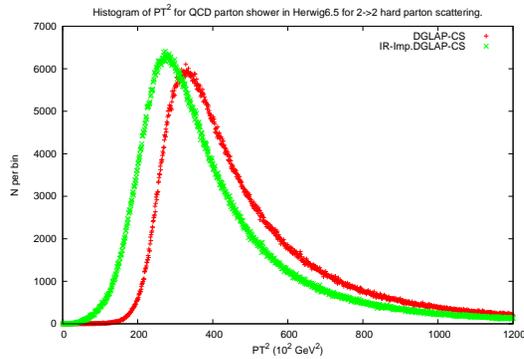


Fig. 10: The  $p_T$ -distribution shower comparison in HERWIG6.5 – preliminary results.

- [2] G. Sterman, Nucl. Phys. **281**, 310 (1987).
- [3] S. Catani and L. Trentadue, Nucl. Phys. **B 327**, 323 (1989).
- [4] G. S. A. Banfi and G. Zanderighi, JHEP **0503**, 073 (2005).
- [5] G. C. et al., JHEP **0101**, 010 (2001).
- [6] S. M. T. Sjostrand and P. Skands, JHEP **0605**, 026 (2006).
- [7] G. Marchesini and B. Webber, Nucl. Phys. **B 238**, 1 (1984).
- [8] G. Marchesini and B. Webber, Nucl. Phys. **B 310**, 461 (1988).
- [9] T. Sjostrand and P. Skands, Eur. Phys. J. **C 39**, 129 (2005).
- [10] G. M. S. Catani and B. Webber, Nucl. Phys. **B 349**, 635 (1991).

- [11] M. Dasgupta and G. Salam, Phys. Lett. **B 512**, 323 (2001).
- [12] M. Dasgupta and G. Salam, JHEP **0208**, 017 (2002).
- [13] G. C. A. Banfi and M. Dasgupta, JHEP **0703**, 050 (2007).
- [14] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C33**, 477 (2004),  
arXiv:hep-ex/0310019.
- [15] D0 Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 221801 (2005),  
arXiv:hep-ex/0409040.
- [16] F. Hautmann and H. Jung, JHEP **10**, 113 (2008), arXiv:0805.1049 [hep-ph].
- [17] M. Dasgupta and G. P. Salam, Phys. Lett. **B512**, 323 (2001),  
arXiv:hep-ph/0104277.
- [18] M. Dasgupta and G. P. Salam, JHEP **03**, 017 (2002), arXiv:hep-ph/0203009.
- [19] N. Kidonakis, G. Oderda, and G. Sterman, Nucl. Phys. **B525**, 299 (1998),  
arXiv:hep-ph/9801268.
- [20] D. de Florian and W. Vogelsang, Phys. Rev. **D76**, 074031 (2007), arXiv:0704.1677  
[hep-ph].
- [21] A. Banfi, M. Dasgupta, and Y. Delenda, Phys. Lett. **B665**, 86 (2008),  
arXiv:0804.3786 [hep-ph].
- [22] A. Banfi, G. Marchesini, Y. L. Dokshitzer, and G. Zanderighi, JHEP **07**, 002 (2000),  
arXiv:hep-ph/0004027.
- [23] G. Dissertori *et al.*, JHEP **02**, 040 (2008), arXiv:0712.0327 [hep-ph].
- [24] A. Gehrmann-De Ridder, T. Gehrmann, E. W. N. Glover, and G. Heinrich, JHEP  
**12**, 094 (2007), arXiv:0711.4711 [hep-ph].
- [25] R. W. L. Jones, M. Ford, G. P. Salam, H. Stenzel, and D. Wicke, JHEP **12**, 007 (2003),  
arXiv:hep-ph/0312016.
- [26] S. Catani, L. Trentadue, G. Turnock, and B. R. Webber, Nucl. Phys. **B407**, 3 (1993).
- [27] R. K. Ellis, D. A. Ross, and A. E. Terrano, Nucl. Phys. **B178**, 421 (1981).
- [28] Z. Kunszt and P. Nason, CERN Yellow Report **1**, 373 (1989).
- [29] W. T. Giele and E. W. N. Glover, Phys. Rev. **D46**, 1980 (1992).
- [30] S. Catani and M. H. Seymour, Phys. Lett. **B378**, 287 (1996),  
arXiv:hep-ph/9602277.
- [31] S. Catani, G. Turnock, B. R. Webber, and L. Trentadue, Phys. Lett. **B263**, 491 (1991).

- [32] S. Catani, G. Turnock, and B. R. Webber, *Phys. Lett.* **B295**, 269 (1992).
- [33] Y. L. Dokshitzer, A. Lucenti, G. Marchesini, and G. P. Salam, *JHEP* **01**, 011 (1998), [arXiv:hep-ph/9801324](#).
- [34] S. Catani and B. R. Webber, *Phys. Lett.* **B427**, 377 (1998), [arXiv:hep-ph/9801350](#).
- [35] E. Gardi and L. Magnea, *JHEP* **08**, 030 (2003), [arXiv:hep-ph/0306094](#).
- [36] A. Banfi, G. P. Salam, and G. Zanderighi, *JHEP* **01**, 018 (2002), [arXiv:hep-ph/0112156](#).
- [37] T. Gehrmann, G. Luisoni, and H. Stenzel, *Phys. Lett.* **B664**, 265 (2008), [arXiv:0803.0695 \[hep-ph\]](#).
- [38] T. Becher and M. D. Schwartz, *JHEP* **07**, 034 (2008), [arXiv:0803.0342 \[hep-ph\]](#).
- [39] ALEPH Collaboration, A. Heister *et al.*, *Eur. Phys. J.* **C35**, 457 (2004).
- [40] M. Dittmar, F. Pauss, and D. Zurcher, *Phys. Rev.* **D56**, 7284 (1997), [arXiv:hep-ex/9705004](#).
- [41] LHC Collaboration, M. Rijssenbeek. Prepared for 14th Topical Conference on Hadron Collider Physics (HCP 2002), Karlsruhe, Germany, 29 Sep - 4 Oct 2002.
- [42] M. Dittmar. Prepared for 14th Topical Conference on Hadron Collider Physics (HCP 2002), Karlsruhe, Germany, 29 Sep - 4 Oct 2002.
- [43] C. Glosser, S. Jadach, B. F. L. Ward, and S. A. Yost, *Mod. Phys. Lett.* **A19**, 2113 (2004), [arXiv:hep-ph/0404087](#).
- [44] B. F. L. Ward, C. Glosser, S. Jadach, and S. A. Yost, *Int. J. Mod. Phys.* **A20**, 3735 (2005), [arXiv:hep-ph/0411047](#).
- [45] B. F. L. Ward, C. Glosser, S. Jadach, and S. A. Yost (2004), [arXiv:hep-ph/0410277](#).
- [46] B. F. L. Ward and S. A. Yost (2005), [arXiv:hep-ph/0509003](#).
- [47] B. F. L. Ward and S. A. Yost (2006), [arXiv:hep-ph/0610230](#).
- [48] B. F. L. Ward and S. A. Yost, *Acta Phys. Polon.* **B38**, 2395 (2007), [arXiv:0704.0294 \[hep-ph\]](#).
- [49] B. F. L. Ward and S. A. Yost, *PoS RADCOR2007*, 038 (2007), [arXiv:0802.0724 \[hep-ph\]](#).
- [50] S. Haywood *et al.* (1999), [arXiv:hep-ph/0003275](#).
- [51] H. Spiesberger, *Phys. Rev.* **D52**, 4936 (1995), [arXiv:hep-ph/9412286](#).

- [52] M. Roth and S. Weinzierl, Phys. Lett. **B590**, 190 (2004), arXiv:hep-ph/0403200.
- [53] J. Blumlein and H. Kawamura, Nucl. Phys. **B708**, 467 (2005), arXiv:hep-ph/0409289.
- [54] W. J. Stirling, A. D. Martin, R. G. Roberts, and R. S. Thorne. Prepared for 32nd International Conference on High-Energy Physics (ICHEP 04), Beijing, China, 16-22 Aug 2004.
- [55] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C39**, 155 (2005), arXiv:hep-ph/0411040.
- [56] M. Ciccolini, A. Denner, and S. Dittmaier, PoS **RADCOR2007**, 014 (2007), arXiv:0712.2895 [hep-ph].
- [57] A. Denner, M. Melles, and S. Pozzorini, Nucl. Phys. **B662**, 299 (2003), arXiv:hep-ph/0301241.
- [58] B. Ward, Advances in High Energy Physics, p. 682312 (2008). Article ID 682312.
- [59] C. Di'Lieto, S. Gendron, I. G. Halliday, and C. T. Sachrajda, Nucl. Phys. **B183**, 223 (1981).
- [60] R. Doria, J. Frenkel, and J. C. Taylor, Nucl. Phys. **B168**, 93 (1980).
- [61] S. Catani, M. Ciafaloni, and G. Marchesini, Nucl. Phys. **B264**, 588 (1986).
- [62] S. Catani, Z. Phys. **C37**, 357 (1988).
- [63] Particle Data Group Collaboration, W. M. Yao *et al.*, J. Phys. **G33**, 1 (2006).
- [64] B. F. L. Ward (2007), arXiv:0707.2101 [hep-ph].
- [65] G. Sterman, Nucl. Phys. **B281**, 310 (1987).
- [66] S. Catani and L. Trentadue, Nucl. Phys. **B327**, 323 (1989).
- [67] S. Catani and L. Trentadue, Nucl. Phys. **B353**, 183 (1991).
- [68] C. W. Bauer, A. V. Manohar, and M. B. Wise, Phys. Rev. Lett. **91**, 122001 (2003), arXiv:hep-ph/0212255.
- [69] C. W. Bauer, C. Lee, A. V. Manohar, and M. B. Wise, Phys. Rev. **D70**, 034014 (2004), arXiv:hep-ph/0309278.
- [70] G. Corcella *et al.* (2002), arXiv:hep-ph/0210213.
- [71] T. Sjostrand, L. Lonnblad, S. Mrenna, and P. Skands (2003), arXiv:hep-ph/0308153.
- [72] S. Frixione and B. R. Webber, JHEP **06**, 029 (2002), hep-ph/0204244.

- [73] S. Frixione, P. Nason, and B. R. Webber, *JHEP* **08**, 007 (2003), [hep-ph/0305252](#).
- [74] N. E. Adam, V. Halyo, and S. A. Yost, *JHEP* **05**, 062 (2008), [arXiv:0802.3251](#) [[hep-ph](#)].
- [75] W. K. Tung *et al.*, *JHEP* **02**, 053 (2007), [arXiv:hep-ph/0611254](#).
- [76] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Phys. Lett.* **B652**, 292 (2007), [arXiv:0706.0459](#) [[hep-ph](#)].
- [77] K. Melnikov and F. Petriello, *Phys. Rev. Lett.* **96**, 231803 (2006), [arXiv:hep-ph/0603182](#).
- [78] K. Melnikov and F. Petriello, *Phys. Rev.* **D74**, 114017 (2006), [arXiv:hep-ph/0609070](#).
- [79] Q.-H. Cao and C. P. Yuan, *Phys. Rev. Lett.* **93**, 042001 (2004), [arXiv:hep-ph/0401026](#).
- [80] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani, *Phys. Rev.* **D69**, 037301 (2004), [arXiv:hep-ph/0303102](#).
- [81] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and M. Treccani, *JHEP* **05**, 019 (2005), [arXiv:hep-ph/0502218](#).
- [82] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini, *JHEP* **12**, 016 (2006), [arXiv:hep-ph/0609170](#).
- [83] C. M. Carloni Calame, G. Montagna, O. Nicrosini, and A. Vicini, *JHEP* **10**, 109 (2007), [arXiv:0710.1722](#) [[hep-ph](#)].
- [84] E. Barberio, B. van Eijk, and Z. Was, *Comput. Phys. Commun.* **66**, 115 (1991).
- [85] E. Barberio and Z. Was, *Comput. Phys. Commun.* **79**, 291 (1994).
- [86] P. Golonka and Z. Was, *Eur. Phys. J.* **C45**, 97 (2006), [arXiv:hep-ph/0506026](#).
- [87] N. E. Adam, V. Halyo, S. A. Yost, and W. Zhu (2008), [arXiv:0808.0758](#) [[hep-ph](#)].
- [88] D. R. Yennie, S. C. Frautschi, and H. Suura, *Ann. Phys.* **13**, 379 (1961).
- [89] K. T. Mahanthappa, *Phys. Rev.* **126**, 329 (1962).
- [90] S. Jadach and B. F. L. Ward, *Comput. Phys. Commun.* **56**, 351 (1990).
- [91] S. Jadach, W. Placzek, E. Richter-Was, B. F. L. Ward, and Z. Was, *Comput. Phys. Commun.* **102**, 229 (1997).
- [92] S. Jadach and B. F. L. Ward, *Phys. Lett.* **B274**, 470 (1992).

- [93] S. Jadach, W. Placzek, and B. F. L. Ward, Phys. Lett. **B390**, 298 (1997), arXiv:hep-ph/9608412.
- [94] S. Jadach, M. Skrzypek, and B. F. L. Ward, Phys. Rev. **D55**, 1206 (1997).
- [95] S. Jadach, W. Placzek, and B. F. L. Ward, Phys. Rev. **D56**, 6939 (1997), arXiv:hep-ph/9705430.
- [96] S. Jadach, B. F. L. Ward, and Z. Was, Phys. Rev. **D63**, 113009 (2001), arXiv:hep-ph/0006359.
- [97] S. Jadach, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **130**, 260 (2000), arXiv:hep-ph/9912214.
- [98] S. Jadach, W. Placzek, M. Skrzypek, B. F. L. Ward, and Z. Was, Comput. Phys. Commun. **140**, 432 (2001), arXiv:hep-ph/0103163.
- [99] B. F. L. Ward and S. Jadach, Acta Phys. Polon. **B33**, 1543 (2002), arXiv:hep-ph/0205062.
- [100] B. F. L. Ward and S. Jadach (2002), arXiv:hep-ph/0210357.
- [101] B. F. L. Ward and S. Jadach, Mod. Phys. Lett. **A14**, 491 (1999).
- [102] D. DeLaney *et al.*, Mod. Phys. Lett. **A12**, 2425 (1997), arXiv:hep-ph/9506302.
- [103] D. B. DeLaney, S. Jadach, C. Shio, G. Siopsis, and B. F. L. Ward, Phys. Rev. **D52**, 108 (1995), arXiv:hep-ph/9403204.
- [104] D. B. DeLaney, S. Jadach, C. Shio, G. Siopsis, and B. F. L. Ward, Phys. Lett. **B342**, 239 (1995), arXiv:hep-ph/9304320.
- [105] C. Lee and G. Sterman, Phys. Rev. **D75**, 014022 (2007), arXiv:hep-ph/0611061.
- [106] S. Mert Aybat, L. J. Dixon, and G. Sterman, Phys. Rev. **D74**, 074004 (2006), arXiv:hep-ph/0607309.
- [107] B. F. L. Ward, Annals Phys. **323**, 2147 (2008), arXiv:0707.3424 [hep-ph].
- [108] E. G. Floratos, D. A. Ross, and C. T. Sachrajda, Nucl. Phys. **B129**, 66 (1977).
- [109] E. G. Floratos, D. A. Ross, and C. T. Sachrajda, Nucl. Phys. **B152**, 493 (1979).
- [110] A. Gonzalez-Arroyo, C. Lopez, and F. J. Yndurain, Nucl. Phys. **B153**, 161 (1979).
- [111] A. Gonzalez-Arroyo and C. Lopez, Nucl. Phys. **B166**, 429 (1980).
- [112] G. Curci, W. Furmanski, and R. Petronzio, Nucl. Phys. **B175**, 27 (1980).
- [113] W. Furmanski and R. Petronzio, Phys. Lett. **B97**, 437 (1980).

- [114] E. G. Floratos, C. Kounnas, and R. Lacaze, Nucl. Phys. **B192**, 417 (1981).
- [115] R. Hamberg and W. L. van Neerven, Nucl. Phys. **B379**, 143 (1992).
- [116] S. Moch, J. A. M. Vermaseren, and A. Vogt, Nucl. Phys. **B688**, 101 (2004), arXiv:hep-ph/0403192.
- [117] A. Vogt, S. Moch, and J. A. M. Vermaseren, Nucl. Phys. **B691**, 129 (2004), arXiv:hep-ph/0404111.
- [118] T. Carli, G. P. Salam, and F. Siegert (2005), arXiv:hep-ph/0510324.
- [119] U. Baur, S. Keller, and W. K. Sakumoto, Phys. Rev. **D57**, 199 (1998), arXiv:hep-ph/9707301.
- [120] U. Baur, S. Keller, and D. Wackerroth, Phys. Rev. **D59**, 013002 (1999), arXiv:hep-ph/9807417.
- [121] U. Baur, O. Brein, W. Hollik, C. Schappacher, and D. Wackerroth, Phys. Rev. **D65**, 033007 (2002), arXiv:hep-ph/0108274.
- [122] S. Dittmaier and M. Kramer, Phys. Rev. **D65**, 073007 (2002), arXiv:hep-ph/0109062.
- [123] V. A. Zykunov, Eur. Phys. J. direct **C3**, 9 (2001), arXiv:hep-ph/0107059.
- [124] R. Hamberg, W. L. van Neerven, and T. Matsuura, Nucl. Phys. **B359**, 343 (1991).
- [125] W. L. van Neerven and E. B. Zijlstra, Nucl. Phys. **B382**, 11 (1992).
- [126] C. Anastasiou, L. J. Dixon, K. Melnikov, and F. Petriello, Phys. Rev. **D69**, 094008 (2004), arXiv:hep-ph/0312266.
- [127] G. Balossini, C. M. Carloni Calame, G. Montagna, O. Nicrosini, and F. Piccinini, Nucl. Phys. Proc. Suppl. **162**, 59 (2006), arXiv:hep-ph/0610022.
- [128] G. Balossini, C. M. Carloni Calame, G. Montagna, O. Nicrosini, and F. Piccinini (2006), arXiv:hep-ph/0609313.
- [129] S. Jadach and M. Skrzypek, Comput. Phys. Commun. **175**, 511 (2006), arXiv:hep-ph/0504263.
- [130] P. Stephens, S. Jadach, and M. Skrzypek, Acta Phys. Polon. **B38**, 2379 (2007).
- [131] T. D. Lee and M. Nauenberg, Phys. Rev. **133**, B1549 (1964).
- [132] B. W. et al., *preprint in pre.* (unpublished). To be published.
- [133] M. Y. Kalmykov, B. F. L. Ward, and S. Yost, JHEP **02**, 040 (2007), arXiv:hep-th/0612240.

- [134] M. Y. Kalmykov, B. F. L. Ward, and S. A. Yost, JHEP **10**, 048 (2007), arXiv:0707.3654 [hep-th].
- [135] M. Y. Kalmykov, B. F. L. Ward, and S. A. Yost, JHEP **11**, 009 (2007), arXiv:0708.0803 [hep-th].
- [136] S. A. Yost, M. Y. Kalmykov, and B. F. L. Ward (2008), arXiv:0808.2605 [hep-ph].
- [137] M. Y. Kalmykov and B. A. Kniehl (2008), arXiv:0807.0567 [hep-th].

# Jets and jet algorithms

*Victor Coco, Pierre-Antoine Delsart, Juan Rojo, Christian Sander Grégory Soyez*

Jets are an important tool in hadronic physics and they will play a predominant role at the LHC. By defining jets as clusters of particles one aims at accessing, from the final-state particles, the underlying hard parton-level processes. Therefore jets are an essential tool for a variety of studies, such as top reconstruction, mass measurements, Higgs and new physics (NP) searches. Furthermore, they are instrumental for QCD studies, *e.g.* for inclusive jet measurements, which in turn constitute an important input for the determination of parton distribution functions. By clustering particles into jets, jet algorithms reduce complicated multiparticle events in simple final states with few jets. This procedure and the way particles are recombined together (*e.g.* the  $E$ - or  $P$ -scheme) is fundamentally non-unique.

In the following we will present recent progress in the description of jets, both from the phenomenological and the experimental points of view. In particular, we will focus on different aspects of the SIScone and anti- $k_t$  jet algorithms. We will also describe jet finding strategies and jet reconstruction and calibration techniques being developed by the LHC experiments ATLAS, CMS and LHCb.

Finally, a recurring question in jet studies is what the best jet definition for a given physics analysis is. We will present a proposal of a characterization of jet-finding “quality” designed to be simple, robust, physical and reasonably representative of common analysis tasks.

## 1 The SIScone and anti- $k_t$ jet algorithms

*Author: Grégory Soyez*

Two broad classes of jet definitions exist. The first one works by defining a distance between pairs of particles, performing successive recombinations of the pair of closest particles and stopping when all resulting objects are too far apart. Algorithms within this clustering class differ by the definition of the distance, frequent choices being  $d_{ij}^2 = \min(k_{t,i}^2, k_{t,j}^2)(\Delta y_{ij}^2 + \Delta\phi_{ij}^2)$  for the  $k_t$  algorithm [1,2], and  $d_{ij}^2 = (\Delta y_{ij}^2 + \Delta\phi_{ij}^2)$  for the Cambridge-Aachen algorithm [3,4].

Cone algorithms make up the second class, where jets are defined as dominant directions of energy flow. One introduces the concept of *stable cone* as a circle of fixed radius  $R$  in the  $y - \phi$  plane such that the sum of all the momenta of the particles within the cone points in the same direction as the centre of the circle. Cone algorithms attempt to identify all the stable cones. Most implementations use a seeded approach to do so: starting from a given seed *e.g.*, a given direction for the centre of the cone, one computes the contents of the cone, takes the resulting momentum as a new direction and iterates until the cone is found stable. The set of seeds can be taken as the set of initial particles (sometimes over a  $p_t$  threshold) or as the midpoints between previously-found stable cones. As we shall see, this iterative method fails to identify *all* stable cones, leading to infrared (IR) or collinear unsafety in the perturbative computations.

Cone algorithms can be split into two sub-classes according to how they deal with the fact that stable cones may overlap. On the one hand, cone algorithms with split-merge identify the

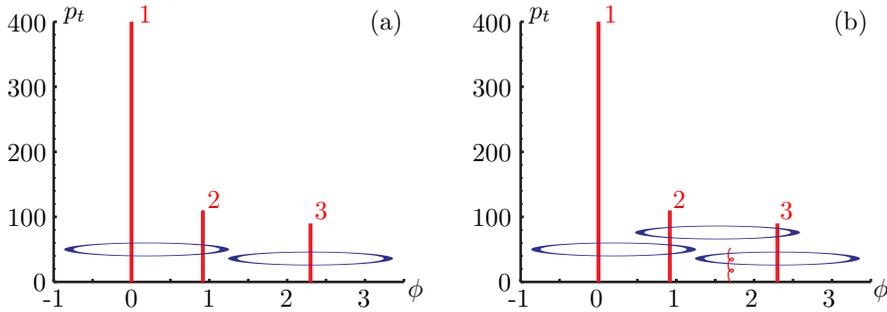


Fig. 1: Stable cones found by the midpoint algorithm for a 3-particle event (left) and for the same event with an additional infinitely soft gluon (right).

hardest overlapping pair of stable cones and merge (split) them if they share more (less) than a fraction  $f$  of the hardest cone. JetClu, midpoint and the ATLAS cone algorithms are typical representatives of that sub-class. On the other hand, cone algorithms with progressive removal start with the hardest unclustered particle, iterate from there until a stable cone is found and call it a jet. Its contents are removed and one starts again with the remaining particles. The CMS iterative cone is the typical example of this second sub-class, with the particular feature that hard jets are fully conical.

The Snowmass accords have established a series of requirements that any jet algorithm has to fulfill. These are basically that one can use the algorithm for theoretical computations, *e.g.* it gives finite perturbative results, as well as for experimental purposes, *e.g.* it runs fast enough and has small corrections from hadronisation and the underlying event.

We show in these proceedings that both the cone algorithms with split-merge and with progressive removal fail to give finite perturbative results. More precisely, we illustrate that midpoint suffers from IR unsafety and the iterative cone is collinear unsafe. We introduce SIScone and the anti- $k_t$  algorithms as infrared- and collinear-safe solutions to those problems that do not spoil the experimental usability. We conclude by discussing the importance of using these new algorithms if we want to take full advantage of jet studies at the LHC.

### 1.1 SIScone as a replacement for the midpoint algorithm

Let us consider the 3-particle event displayed in Fig. 1(a). When clustered with the midpoint algorithm, 2 stable cones are found, leading to two jets: one with particles 1 and 2 and a second one with particle 3. If one adds to that hard event an infinitely soft gluon as shown in Fig. 1(b), a third stable cone is found and the three hard particles are clustered in a single jet. This change in the jet structure upon addition of soft particles, a phenomenon which happens with infinite probability in perturbative QCD, gives rise to divergences in the perturbative expansion and proves that the midpoint algorithm is IR unsafe<sup>1</sup>. Note also that the situation is even worse with JetClu or the ATLAS cone algorithms, where the IR unsafety is already present in events with 2 particles, *i.e.* one order earlier in the perturbative expansion.

<sup>1</sup>Note that when a seed threshold is used, the midpoint algorithm becomes collinear unsafe.

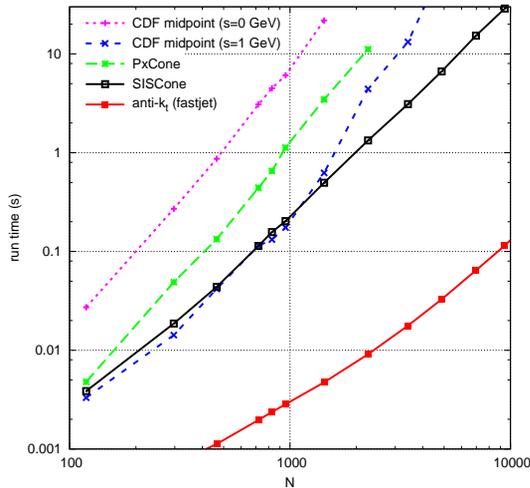


Fig. 2: Clustering time for SISConc compared to typical implementations of the midpoint and anti- $k_t$  algorithms.

This problem arises from the fact that the seeded approach misses stable cones — here the one containing particles 2 and 3 in Fig. 1(a). The workaround to restore IR safety is thus to find a seedless method that provably identifies all stable cones. This is notoriously complex: a naive approach testing the stability of all subsets of particles [5] has a complexity of order  $N \times 2^N$  for  $N$  particles which is much slower than the  $\mathcal{O}(N^3)$  complexity of the midpoint algorithm, making this solution unusable for experimental purposes.

The solution [6] is to use the geometrical observation that any enclosure in the  $y - \phi$  plane can be moved without changing its contents until it touches two points. Browsing all pairs of particles allows thus to enumerate all possible cones and to check their stability at an overall cost of  $\mathcal{O}(N^3)$ . Additional efforts to limit the amount of full stability tests to its minimum can even bring the final complexity to  $\mathcal{O}(N^2 \log(N))$ , *i.e.* faster than the midpoint algorithm. This has been implemented [6–9] in a C++ code named SISConc (Seedless Infrared Safe Cone). Fig. 2 illustrates the fact that in practice SISConc runs faster than the typical implementations of the midpoint algorithm without a seed threshold and at least as fast as when a 1 GeV seed threshold is used.

Therefore, SISConc is the first cone algorithm to satisfy the Snowmass requirements, that is to be at the same time IR and collinear safe, and to be fast enough to be used in experimental analysis.

## 1.2 Anti- $k_t$ as a replacement for the iterative cone algorithm

As for the midpoint algorithm, we start by considering an event with three hard particles (see Fig. 3(a)). When clustered with the iterative cone, iteration starts with particle 2, one stable cone containing all particles is found, resulting in a 1-jet event. If we now split the hardest particle (2) into two collinear particles (2a and 2b) — a process that also happens with an infinite probability in perturbative QCD — as shown on Fig. 3(b), clustering with the iterative cone now starts with

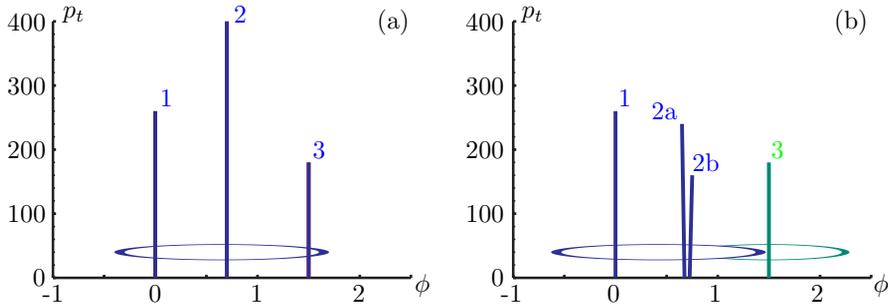


Fig. 3: Jets found by the iterative cone for a 3-particle event (left) and for the same event with a collinear splitting (right).

particle 1 which, after iteration, gives a first jet made of particle 1 plus the two collinear ones, then a second jet with particle 3. This example proves that the iterative cone algorithm is collinear unsafe.

Quite surprisingly, we can find a solution to that problem by coming back to the class of the recombination algorithms. The distance measures introduced earlier can be written as

$$d_{ij}^2 = \min(k_{t,i}^{2p}, k_{t,j}^{2p})(\Delta y_{ij}^2 + \Delta\phi_{ij}^2),$$

with  $p = 1$  for the  $k_t$  algorithm and  $p = 0$  for the Cambridge-Aachen algorithm. We can then consider a third case, the one for which  $p = -1$  and call it the *anti- $k_t$  algorithm* [10]. Obviously, this algorithm is IR and collinear safe. Furthermore, its implementation can benefit from the same geometrical observations that allowed for fast implementation of the  $k_t$  algorithm [8]. The anti- $k_t$  algorithm thus runs at a speed similar to the one of the  $k_t$  algorithm, which certainly makes it usable for experimental purposes as seen on Fig. 2.

To understand the link between the anti- $k_t$  algorithm and the iterative cone algorithm, we note from the definition of the anti- $k_t$  distance that pairs involving a hard particle will be given small distances. This means that soft particles will be recombined with hard ones before recombining among themselves. As a result, the hard jets will have a circular boundary. This soft-resilience of the anti- $k_t$  algorithm is exactly the hallmark of the iterative cone and it is in that respect that the anti- $k_t$  can be seen as its IR and collinear safe replacement.

To illustrate this property, we show in Fig. 4 the jets resulting from the clustering of an event made with a few hard particles and a large number of very soft ones uniformly dis-

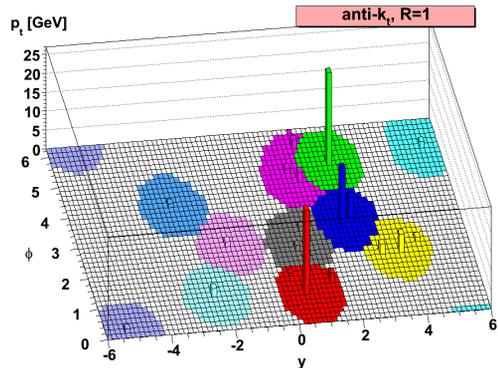
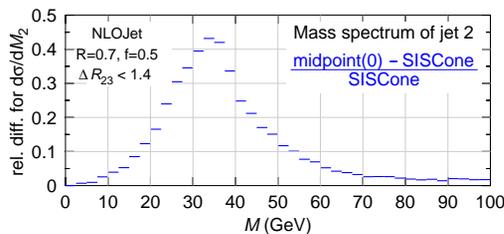


Fig. 4: Illustration of the regularity of the jets obtained with the anti- $k_t$  algorithm.

Observable	first miss cones at	Last meaningful order
Inclusive jet cross section	NNLO	NLO
$W/Z/H + 1$ jet cross section	NNLO	NLO
3 jet cross section	NLO	LO (NLO in NLOJet)
$W/Z/H + 2$ jet cross sect.	NLO	LO (NLO in MCFM)
jet masses in 3 jets	LO	<b>none</b> (LO in NLOJet)

Table 1: Perturbative level at which IR or collinear unsafety arises for various processes.

Fig. 5: Mass of the 2<sup>nd</sup> hardest jet in 3-jet events: relative difference between midpoint and SIScone. The 2<sup>nd</sup> and 3<sup>rd</sup> jets are imposed to be distant by at most  $2R$ .

tributed on a grid in the  $y - \phi$  plane. It is clear that the hardest jets are perfectly circular and that, in general, the boundaries between the jets are regular.

### 1.3 Physical impact and discussion

As we have seen, the seeded approach to stable cone search suffers from problems with respect to perturbative QCD expansion: the algorithms with split-merge are IR unsafe, while the iterative cone (with progressive removal) is collinear unsafe. We have introduced SIScone as a natural replacement of the cone algorithms with split-merge like midpoint, and the anti- $k_t$  algorithm as a candidate to replace the iterative cone. These new algorithms are both IR and collinear safe.

The question one might ask is to what extent these IR and collinear safety issues are important in real measurements. Since the unsafety arises when one has 3 particles in a common vicinity, it becomes important at the order  $\alpha_s^4$  or  $\alpha_{EW}\alpha_s^3$  of the perturbative series.

Table 1 summarises for different physical processes, the order at which seeded algorithms like midpoint of the iterative cone stop to be valid. The main message we can get from that table is that, if we do not want theoretical efforts in precise QCD computations to be done in vain, the resort of an IR and collinear safe algorithm like SIScone and the anti- $k_t$  is fundamental.

To illustrate the argument more quantitatively, Fig. 5 shows the relative difference, expected to be present at the LO of perturbative QCD, between SIScone and midpoint for the mass of the second hardest jet in 3-jet events. Differences reaching up to 40% are observed, proving that an IR and collinear safe algorithm is mandatory. The situation is even worse with JetClu or the ATLAS cone algorithm. As the infrared-unsafety problem becomes apparent at the order  $\alpha_s^3$  or  $\alpha_{EW}\alpha_s^2$ , *i.e.* one order earlier than with midpoint.

## 2 Quality measures for jet finding at the LHC

Author: *Juan Rojo*

A recurring question in jet studies is what the best jet definition for a given physics analysis is. In this contribution we propose a characterization of jet-finding “quality” designed to be simple, robust, physical and reasonably representative of common analysis tasks.

For this purpose, we require a source of quarks and gluons with well-defined energies. We will obtain these from Monte Carlo production and decay of fictitious narrow  $Z'$  and  $H$  bosons, with  $Z' \rightarrow q\bar{q}$  and  $H \rightarrow gg$  generated with Pythia 6.5 [11] with di-jet invariant masses ranging from 100 GeV to 4 TeV. For each generated event we will cluster the event into jets with about 50 different jet definitions, where a jet definition, JD, consists of the jet algorithm and the associated parameters, like the radius  $R$  [12]. The radius  $R$  will be varied between 0.3 and 1.5. For each event, we determine the invariant mass of the sum of the two hardest jets. The distribution of invariant masses should then have a peak near the heavy boson mass. We will take the sharpness of that peak to be indicative of the quality of each jet definition.

The infrared- and collinear-safe (IRC) safe jet algorithms under scrutiny are the longitudinally invariant inclusive  $k_t$  algorithm [1,2,13], the Cambridge-Aachen (C/A) algorithm [3,4], the anti- $k_t$  algorithm [10], SIScone [6] as well as C/A with filtering. The latter is C/A supplemented with a filtering procedure [14] in which, subsequent to the jet finding, each jet is unclustered down to subjets at angular scale  $x_{\text{filt}}R$  and one retains only the  $n_{\text{filt}}$  hardest of the subjets. We use  $x_{\text{filt}} = 0.5$  and  $n_{\text{filt}} = 2$ . All the jet algorithms have been used in the implementations and/or plug-ins of the `FastJet` package [8], version 2.3, with the exception of C/A with filtering, which will be made public in a forthcoming `FastJet` release.

This contribution summarizes work [15] in collaboration with M. Cacciari, G. Salam and G. Soyez, initiated in the context of the “Les Houches Physics at TeV colliders 2007” workshop [12].

### 2.1 Quality measures and effective luminosity ratio

As described in detail in [15], the merit of the jet finding is quantified by two quality measures:

1.  $Q_{f=z}^w$ : the width of the smallest (reconstructed) mass window that contains a fraction  $f = z$  of the generated massive objects,

$$f \equiv \left( \frac{\# \text{ reco. massive objects in window of width } w}{\text{Total } \# \text{ generated massive objects}} \right) = z. \quad (1)$$

2.  $Q_{w=x\sqrt{M}}^{1/f}$ : to compute this quality measure, we take a window of fixed width  $w$  and slide it over the mass distribution so as to maximise its contents. Then the figure of merit is given by

$$Q_{w=x\sqrt{M}}^{1/f} \equiv \left( \frac{\text{Max } \# \text{ reco. massive objects in window of width } w = x\sqrt{M}}{\text{Total } \# \text{ generated massive objects}} \right)^{-1}, \quad (2)$$

It is clear from its definitions that the smaller the quality measures, the better the corresponding jet definition. An illustrative example of these two measures is shown in Fig. 6. We observe that the quality measures quantify the intuitive assessment of the goodness of jet finding, represented by the sharpness of the reconstructed invariant mass peak. Note that in our approach, any matching to non-physical quantities like Monte Carlo partons is deliberately avoided.

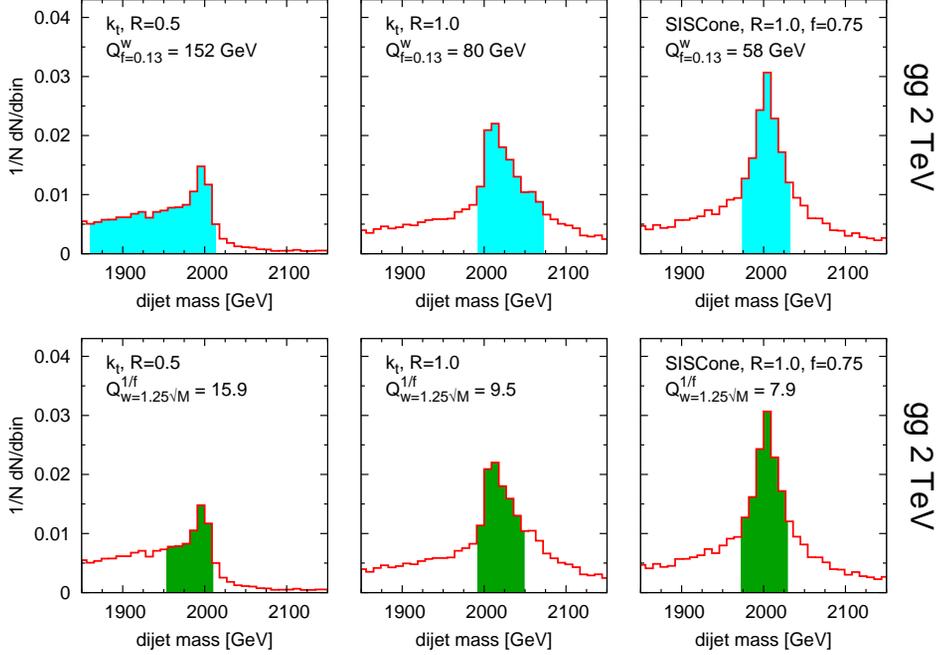


Fig. 6: Di-jet invariant mass distributions for the  $gg$  case at  $M = 2$  TeV, comparing three jet definitions for each process. The shaded bands indicate the region used when obtaining the two different quality measures.

These quality measures can be mapped to the corresponding variation of integrated luminosity needed to maintain constant signal significance. As we have seen, a larger quality measure indicates a worse jet definition. This in turn implies that a larger luminosity will be needed to obtain a given significance. It is convenient to express this in terms of an effective luminosity ratio,

$$\rho_{\mathcal{L}}(\text{JD}_2/\text{JD}_1) \equiv \frac{\mathcal{L}(\text{needed with JD}_2)}{\mathcal{L}(\text{needed with JD}_1)} = \left[ \frac{\Sigma(\text{JD}_1)}{\Sigma(\text{JD}_2)} \right]^2. \quad (3)$$

with the signal significance defined in the usual way  $\Sigma(\text{JD}) \equiv N_{\text{signal}}^{\text{JD}} / \sqrt{N_{\text{bkgd}}^{\text{JD}}}$ . Given a certain signal significance with  $\text{JD}_1$ ,  $\rho_{\mathcal{L}}(\text{JD}_2/\text{JD}_1)$  indicates the factor more luminosity needed to obtain the same significance with  $\text{JD}_2$ . For example, the expression for  $\rho_{\mathcal{L}}$  in terms of the first quality measure is

$$\rho_{\mathcal{L}}(\text{JD}_2/\text{JD}_1) = \frac{Q_{f=z}^w(\text{JD}_2)}{Q_{f=z}^w(\text{JD}_1)}. \quad (4)$$

A non-trivial check of the robustness of our analysis is that the luminosity ratios obtained with the two different quality measures are roughly consistent with each other.

## 2.2 Results

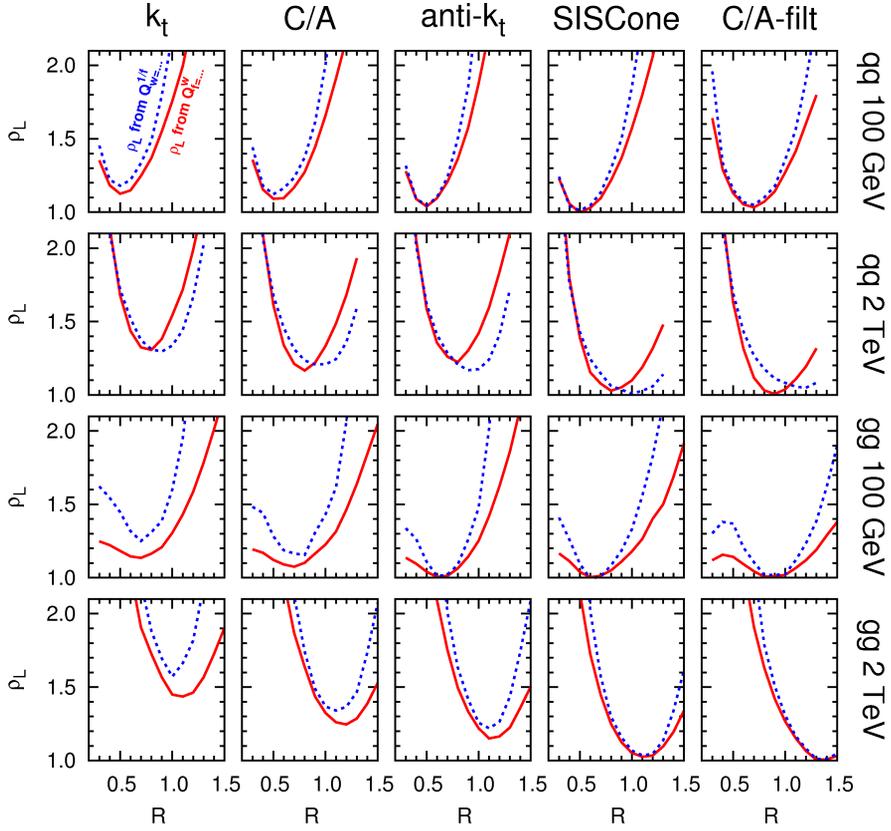


Fig. 7: The effective luminosity ratio, Eq. 3, for quark and gluon jets at 100 GeV and 2 TeV, for all algorithms studied. The two curves in each plot correspond to the value of  $\rho_L$  computed from the respective quality measure. For each process,  $\rho_L$  is normalized to the corresponding optimal jet definition.

Now we present selected results for the effective luminosity ratio for the different cases considered. We show in Fig. 7 a summary of the performance of the various jet definitions studied, for quark and gluon jets at 100 GeV and 2 TeV, without pile-up (PU). First of all, we observe a strong dependence of  $\rho_L$  with respect to  $R$ , as well as sizable differences between jet algorithms. SIScone and C/A-filt turn out to be the optimal jet algorithms in all studied processes. They achieve limited sensitivity to the Underlying Event (UE) while maintaining their perturbative reach. The optimal value of  $R$  grows with the scale of the process, specially for gluon jets, reflecting the interplay between perturbative and non-perturbative effects [16]. Our studies imply that at the TeV scale, rather large values of  $R \sim 1$  are required to obtain optimal

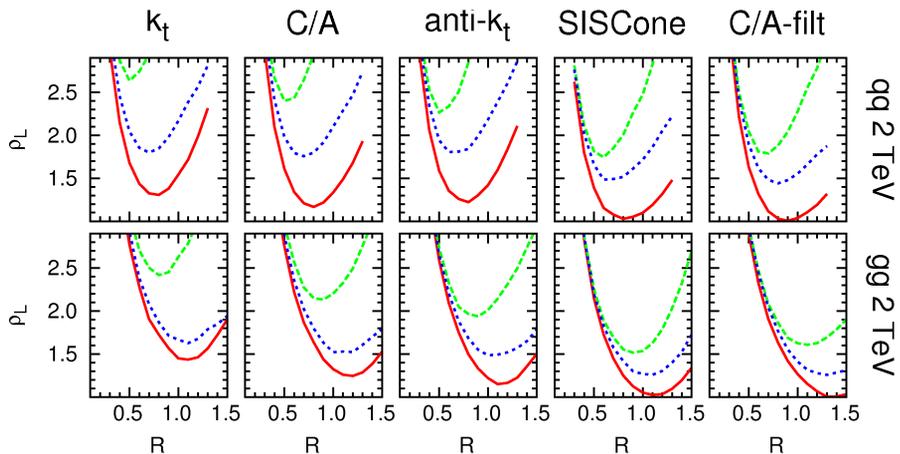


Fig. 8: The effective luminosity ratio, for quark and gluon jets at 100 GeV and 2 TeV, for all five algorithms studied. The red solid lines correspond to the no-PU case, the green dashed lines to the high luminosity PU case while the blue dotted curves correspond to high luminosity with PU subtracted as explained in the text.

resolution. LHC experiments, on the other hand, plan to use smaller radii in general, see for example Ref. [17].

From Fig. 7 one can determine how much more luminosity will be required with a less favoured jet definition compared with the optimal one. For example, we see that for the  $gg$  case at 2 TeV, if the  $k_t$  algorithm is used instead of the optimal one (SISConc), then 50% more luminosity will be required to achieve the same signal significance even at the respective optimal values of  $R$ .

These results are robust against high-luminosity PU [15] once PU is subtracted using the FastJet area method [18, 19], as can be seen in Fig. 8. This has the important consequence that for a given process, a single jet definition could be used at the LHC regardless of the machine luminosity.

As a practical application of our studies, one can consider the impact of less favoured jet definitions in LHC searches with similar signatures. For example, let us consider a particular scenario in which a di-jet invariant mass distribution is reconstructed and let us assume that the jet clustering is performed with a jet definition,  $JD_2$ , whose quality is far from the optimal one,  $JD_1$ , so that the effective luminosity ratio is large, say  $\rho_{\mathcal{L}} \sim 2$ . The net effect of the choice of such non-optimal jet definition for the kinematical reconstruction can be summarized schematically in Fig. 9: the use of  $JD_1$  rather than  $JD_2$  would lead to a discovery signal with approximately only half of the machine running time required with the original jet definition.

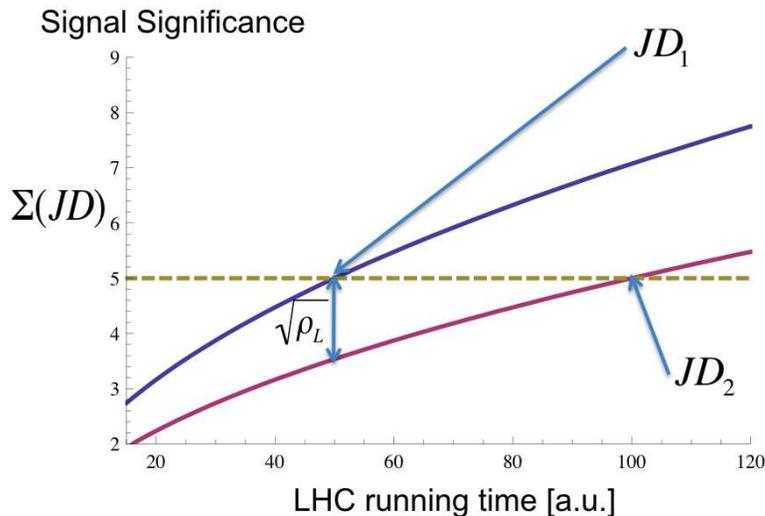


Fig. 9: Example of how optimizing the jet definition might lead to discoveries in less machine running time, compared to the non-optimal one.

### 2.3 Conclusions

Summarizing, we have proposed a technique to quantify the performance of jet algorithms for kinematic reconstructions at the LHC. To allow for more detailed studies of the results of Ref. [15], an interactive webpage has been created at <http://quality.fastjet.fr>, which allows the user to test the effects of changing and modifying various jet definitions and other inputs like PU luminosity for the process under scrutiny.

## 3 Performance of jet reconstruction at CMS

*Author: Christian Sander (on behalf of the CMS Collaboration)*

Almost every process of interest at the LHC contains quarks or gluons in the final state. The partons can not be observed directly, but fragment into stable hadrons, which can be detected in the tracking and calorimeter systems. Calorimeter jets are expected to yield a good description of both the parton-level and the hadron showers emerging from the hard interaction. For Monte Carlo (MC) events, the hadron-level is defined by applying the same clustering algorithms, which are typically formulated to accept any set of four-vectors as input to all stable particles from the MC truth record (“GenJets”). Hadron-level is also referred to as “particle-level”, and jet energy scale corrections based on MC and later on data-driven methods are derived to correct back to this detector independent level. Calorimeter jets are reconstructed using energy deposits in calorimeter towers (“CaloTowers”) as inputs: they are composed of one or more hadronic calorimeter (HCAL) cells and corresponding electromagnetic calorimeter (ECAL) crystals.

The studies presented in what follows are based on QCD di-jet and  $t\bar{t}$  MC samples without pile-up. It is often necessary to associate CaloJets with GenJets in these samples to probe how well the calorimeter-level reconstruction represents the hadron-level of the process. This

association is based on spatial separation in the  $\eta$ - $\phi$ -space between the two jet axes by requiring

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$$

to be less than a certain value. Besides good correspondence to the parton-level and hadron-level, a successful jet algorithm should fulfill two important requirements. Firstly, it should be collinear-safe, such that the outcome remains unchanged if *e.g.* the energy carried by a single particle is instead distributed among two collinear particles. Collinear safety is typically endangered if the jet finding is based on energetic seeds and a threshold is applied to these seeds. Secondly, it should be infrared-safe, such that the result of the jet finding is stable against the addition of soft particles. Jet algorithms which don't comply with either or both of these requirements yield ambiguous results and lead to unnecessary uncertainties when applied to calculations in perturbative theory. The performance of the following four jet clustering algorithms is discussed:

- The Iterative Cone algorithm is a simple seeded cone-based algorithm employed by CMS online in the High Level Trigger (HLT). It has a short and predictable execution time, but is neither collinear- nor infrared-safe.
- The Midpoint Cone [5] algorithm is similar to the Iterative Cone, but infrared-safety is addressed by considering the midpoints between each pair of close (proto-)jets as additional seeds. Despite its improvements to the cone-based clustering procedure, the algorithm has been shown not to be infrared-safe. This algorithm is no longer supported by CMS.
- SIS Cone [6] is the ‘‘Seedless Infrared-Safe Cone’’ jet algorithm. It is collinear- and infrared-safe to all orders of perturbative QCD (pQCD) and demands only slightly higher execution time compared to the Midpoint Cone algorithm.
- fast- $k_T$  [8] is a recent implementation of the  $k_T$  algorithm [1] which is also collinear- and infrared-safe. It has a dramatically reduced execution time with respect to previous implementations of the  $k_T$  algorithm.

The ‘‘ $E$ -Scheme’’ is used for all algorithms as the recombination scheme: the energy and momentum of a jet are defined as the sums of energies and momenta of its constituents. The execution time of the fast- $k_T$  algorithm is comparable to the Iterative Cone algorithm without the discussed deficiencies of the latter. The SIS Cone algorithm requires more CPU resources compared to the Midpoint Cone algorithm. The time spent for the jet reconstruction (0.02 s) of each event however is small compared to the total event reconstruction time (10 s): the particular jet algorithm choice does not impact the overall CPU requirements.

### 3.1 Summary of Jet Performance Study

The performance of the CMS calorimeters is known to be different in the barrel, endcaps and forward regions. Here we focus on the relative performance between different algorithms and radius parameter choices currently supported for CMS analysis. Only distributions for the barrel region are therefore shown. Further details can be found in [20].

The jet matching efficiency is defined as the ratio of the number of particle jets matched to a calorimeter jet within  $\Delta R < 0.5$  to the total number of particle jets. It represents a meaningful measure of the reconstruction efficiency of each jet algorithm, but is strongly correlated

to the position resolution and therefore depends on the  $\Delta R$  cut and the jet size parameter. However, relative comparisons between different algorithms using equivalent size parameters remain instructive. The matching efficiencies for small (left) and large (right) radius parameters as a function of the MC truth  $p_T^{\text{gen}}$  are shown in Fig. 10. The efficiencies of jets reconstructed with the fast- $k_T$  and SIS Cone algorithms indicate better performance than jets reconstructed with the Midpoint Cone and Iterative Cone algorithms.

For the jet response,  $R_{\text{jet}} = p_T/p_T^{\text{gen}}$ , very good agreement between the individual algorithms is found for all regions of the detector, indicating good correspondence between the values of  $D$  for the fast- $k_T$  algorithm and  $R$  for cone algorithms which are being compared [20].

The  $\eta$  resolutions for jets in the barrel region are shown as a function of  $p_T^{\text{gen}}$  in Fig. 11. Good agreement is found among all algorithms with comparable radius parameter, with marginal differences at low  $p_T^{\text{gen}}$ . Jets reconstructed with larger radius parameters yield slightly worse resolution. Note that the position of the primary vertex is assumed to be at  $z = 0$ , which dilutes the  $\eta$  resolution with respect to taking the correct position measured with the tracking detectors into account. The  $\phi$  resolutions can be found in [20].

Fig. 12 shows the jet energy resolutions derived from MC truth for jets in the barrel region. Jets reconstructed with fast- $k_T$  show slightly worse resolution at low  $p_T^{\text{gen}}$ , while no significant impact of the radius parameter choice is observed. The typical jet energy resolution at high  $p_T$ , 100 GeV or 1 TeV, is  $\sim 14\%$  and  $\sim 7\%$  respectively, with no significant dependence on the jet clustering algorithm.

The jet reconstruction performance in  $t\bar{t}$  events is studied by selecting events with one (“lepton+jets”) or zero (“alljets”) electron(s) or muon(s) in the final state from a  $t\bar{t}$  sample with no additional jets (“ $t\bar{t}+0$  jets”).  $t \rightarrow bq\bar{q}'$  and  $\bar{t} \rightarrow \bar{b}\bar{q}q'$  decays are identified on particle level and only events are considered for which all three decay products of one or both  $t(\bar{t})$  decay(s) can be uniquely matched to reconstructed calorimeter jets. The efficiency to select these decays indicates the performance of the respective jet algorithm in a busy multi-jet environment and its ability to correctly resolve the topology of the underlying process. The fast- $k_T$  algorithm is hereby found to fully resolve hadronic  $t(\bar{t})$  decays on calorimeter level more efficiently than any cone-based algorithm. For the selected events, the invariant two-jet ( $W$  boson) and three-jet (top quark) masses are compared on particle-level, calorimeter-level, corrected calorimeter-level, and corrected calorimeter-level with additional flavor-dependent corrections applied. The  $m_W$  and  $m_t$  distributions obtained for all correction levels are shown in Fig. 13 for jets reconstructed with fast- $k_T$   $D = 0.4$ . From the width of the obtained invariant mass distributions one can see that the impact of detector effects on the mass resolution are stronger than the algorithmic differences. A full comparison of the widths of the reconstructed  $m_W$  and  $m_t$  distributions can be found in [20].

### 3.2 Conclusion

The performance comparisons presented include jet energy response, position resolutions, energy resolutions and efficiencies in QCD di-jet samples. We find similar performance at the calorimeter level between algorithms with similar size parameter. The impact of detector effects appears to be more pronounced than the algorithmic differences studied here. The SIS Cone algorithm performs as well as or better than the Midpoint Cone, while known to be preferred

theoretically. Therefore it was decided to adopt SIS Cone as the default cone-based jet algorithm and consequently to include it in the reconstruction in the standard event processing at CMS.

The fast- $k_T$  algorithm is infrared- and collinear safe to all orders of pQCD as well and complementary to the cone-based algorithms. The execution time of fast- $k_T$  is dramatically reduced with respect to earlier implementations and it is therefore well suited for the high multiplicity environment of LHC  $pp$  collisions. We find that it performs as good or better than any other compared algorithm and strongly encourage its use as an alternative to SIS Cone.

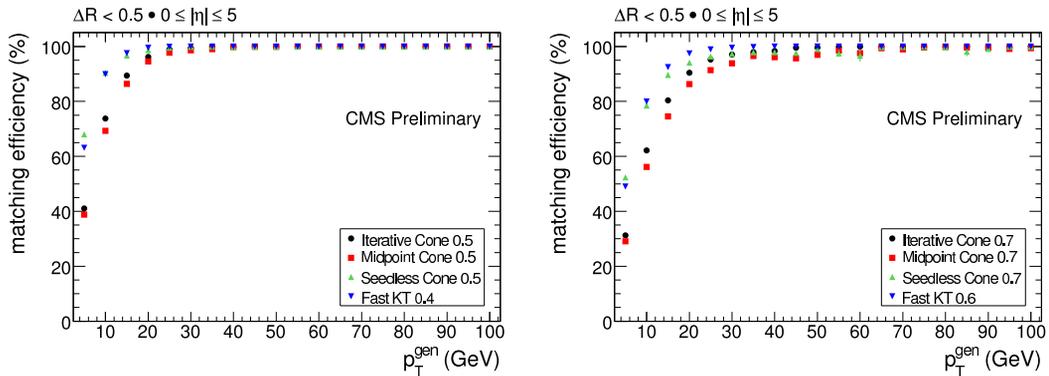


Fig. 10: Matching Efficiency versus  $p_T^{\text{gen}}$  for  $R = 0.5/D = 0.4$  (left) and  $R = 0.7/D = 0.6$  (right) jets.

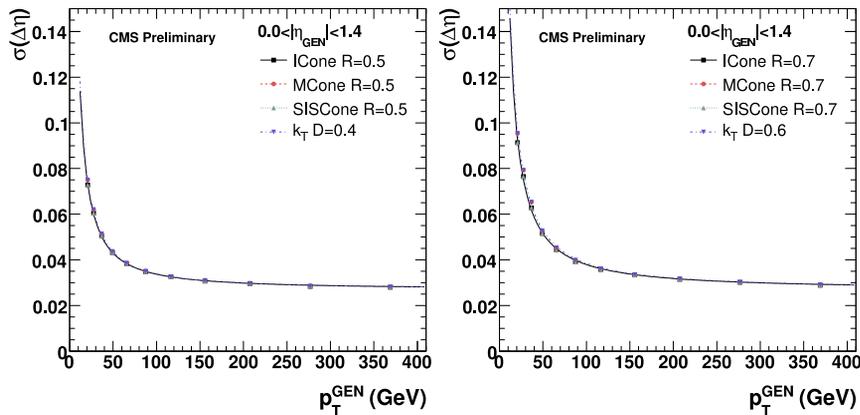


Fig. 11: The jet  $\eta$  resolutions as a function of  $p_T^{\text{gen}}$ , averaged over the Barrel region, for jets clustered with smaller (left) and larger (right) size parameters. The resolutions are derived using MC truth information.

## 4 Jet finding strategies in ATLAS

*Author: Pierre-Antoine Delsart (on behalf of the ATLAS Collaboration)*

ATLAS is a general purpose experiment at the Large Hadron Collider (LHC) [21]. Its calorimetry system, the principal tool for hadronic jet measurements, is described in detail in [21], chapter 5. Some key features of this calorimeter relevant to jet finding are its wide

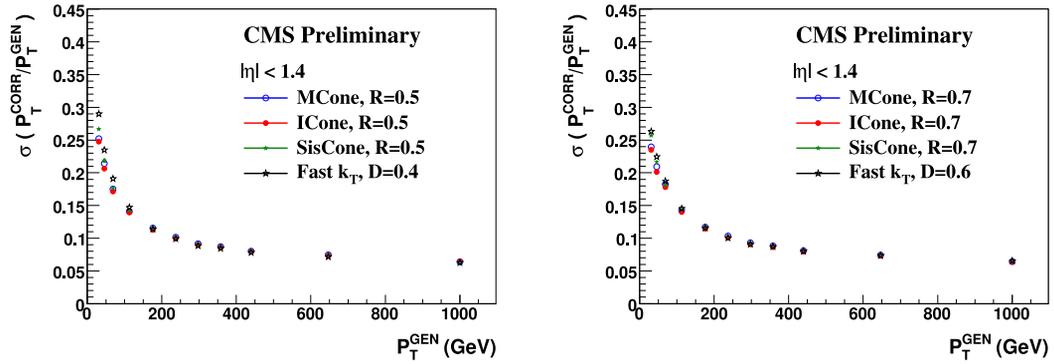


Fig. 12: Jet energy resolution derived from MC truth for Midpoint Cone, Iterative Cone, SIS Cone and fast- $k_T$  with  $R = 0.5/D = 0.4$  (left) and  $R = 0.7/D = 0.6$  (right) in the barrel region ( $|\eta| < 1.4$ ).

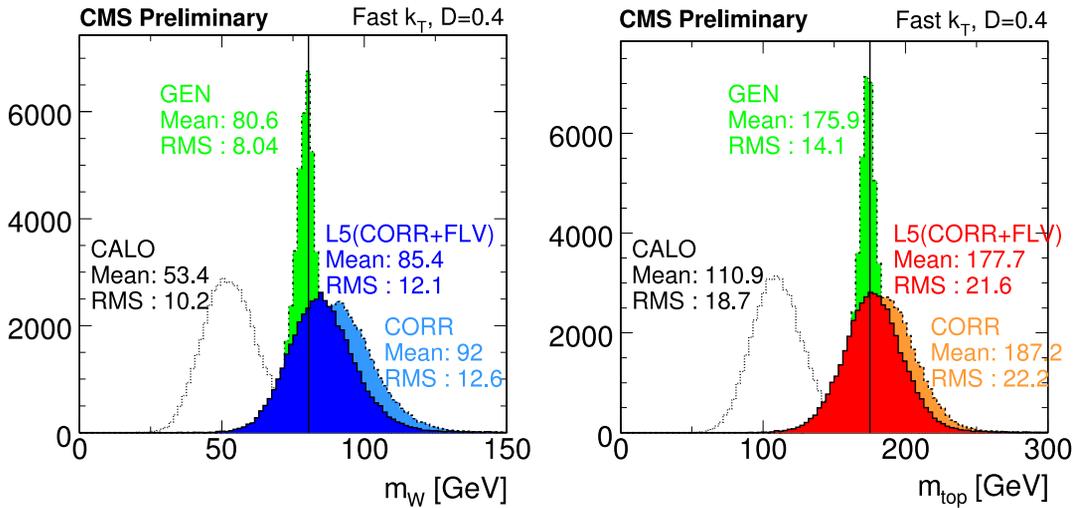


Fig. 13:  $m_W$  and  $m_t$  distributions for hadronic top decays reconstructed with the fast- $k_T$  algorithm,  $D = 0.4$ . Distributions are shown for particle-level jets (GEN), calorimeter jets (CALO), calorimeter jets corrected with “MC-Jet” corrections (CORR), and corrected calorimeter jets with an additional flavour correction applied (L5). Only jets with uncorrected  $p_T \geq 15$  GeV and  $|\eta| \leq 5$  are considered. The generated  $W$  boson (80.42 GeV) and top quark (175 GeV) masses are indicated by the black vertical lines.

acceptance (up to  $|\eta| = 4.9$  in the Forward Calorimeter) and a fine granularity (including up to 7 longitudinal segmentations). On the other hand, the calorimeter is non-compensating (ratio  $1.3 < e/\pi < 1.6$  depending on the specific sub-calorimeter) and this causes the major source of uncertainty in energy measurements because of the large fluctuations of the electromagnetic component of hadronic showers.

The other main experimental challenge will come from the LHC environment : a very large phase space for underlying event, multiple interactions per bunch crossing (23 at full luminosity). Out-of-time pile-up is also expected because of the slow response of the liquid argon calorimeter which will integrate several events before and after a given interaction.

In order to take up these challenges the ATLAS collaboration chose to adopt a flexible approach. In particular, the collaboration is studying two calibration strategies, several in-situ and data-based correction methods, and has designed a software able to cope with any jet algorithm used in physics analysis.

#### 4.1 Jet reconstruction and calibration

The ATLAS jet-related software is designed to allow any input to jet finding algorithms, provided the input is a set of valid four-momenta. This allows to run exactly the same jet finders on Monte Carlo truth simulated particles, real signal, tracks, etc. Two different calorimeter signal definitions are considered as input signal for jet finding:

- Calorimeter towers : all cells in the same projective direction (defined by a grid in the  $(\eta, \phi)$  plane) are grouped into a tower. The four-momentum is formed by the sum of the cells energies, possibly including a geometrical weight for cells larger than the tower grid size, and the direction of the tower.
- Topological clusters (“TopoClusters”). Cells are clustered together in the 3 dimensions of the calorimeter according to a nearest neighbour algorithm [21] which intrinsically performs a noise suppression.

Besides different types of input signal, ATLAS considers two approaches for the jet calibration.

**Global hadronic calibration.** Jets are built from raw calorimeter signal (towers or clusters), then a set of correction factors (weights) are applied to the energy of the constituting cells. The weights depend on the characteristics of the cells, in particular its energy density and its location in the calorimeter. They are extracted from a fit to simulated di-jet events. With this method, all calibration corrections are included in a single set of weights, hence its name “global” [22].

**Local hadronic calibration.** This second method is an attempt to have a finer, better understood calibration method for jets. It relies on hadronic calibration of topological clusters [23] : jets are built from these calibrated input signals. Then a jet energy scale correction remains to be applied. This approach is more complex but allows to decouple different corrections (non-compensation, dead material losses, energy scale) and is thus very promising.

An illustration of performance for both calibration methods is shown in Fig. 14; a detailed discussion of these performances can be found in Ref. [24];

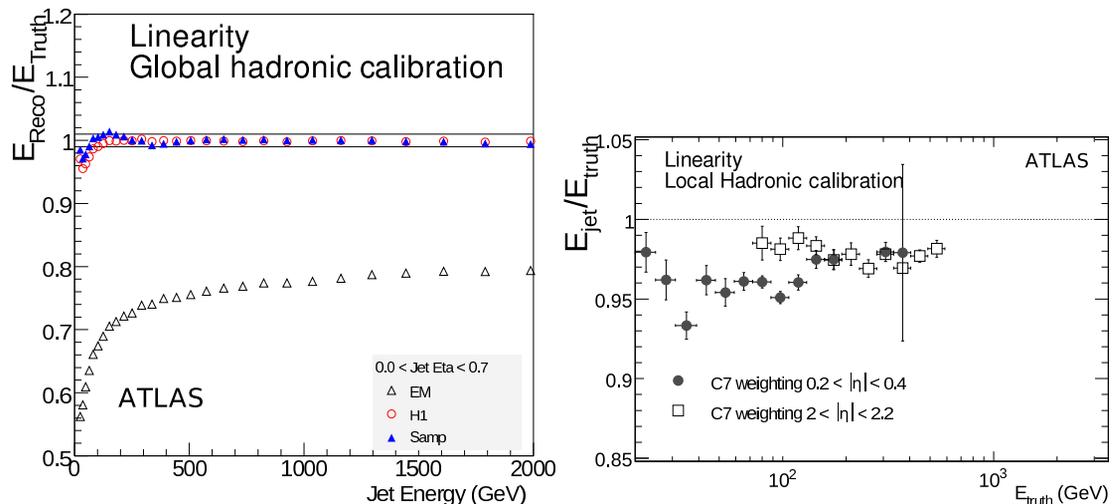


Fig. 14: Left : linearity of jet reconstruction with global calibration (QCD di-jet sample). Red and blue marks correspond to 2 alternative global calibration methods. Right: linearity with local calibration (before energy scale correction).

## 4.2 Data driven corrections

In parallel to the base-line calibrations described above, ATLAS aims to reach a precise energy scale measurement using experimental data directly. Several methods are studied:

- Momentum measurement from the tracker (P) compared to energy deposition in calorimeter (E) allows to validate the energy scale for charged pions in minimum bias events [25], studying the E/P ratio.
- QCD di-jet events can be used to uniformize the response of the calorimeter in  $\eta$  and  $\phi$  [26].
- Z+jets or  $\gamma$ +jets events will be used assuming an excellent calibration of electromagnetic objects : applying  $P_T$  balance or missing  $E_T$  projection techniques will allow to retrieve a correct jet energy scale [26].
- With QCD multi-jet events it is possible to correct high- $p_T$  jets against several lower  $p_T$  jets whose energy scale is better known thanks to previous methods [26].

In some analyses, *in-situ* methods using constraints coming from the mass of the W boson will be applied to control even better the jet energy. This is typically the case in top physics analyses where different methods are under study [27].

## 4.3 Jets algorithms, other jets studies

Several different jet algorithms are available for physics analysis. Two families of such algorithms are reconstructed by default :

- ATLAS iterative cone algorithm (described in detail in [22]), with cone radii 0.4 and 0.7;
- Kt clustering algorithm [22], setting the size parameter D to 0.3 and 0.6 .

These algorithms are officially supported and used in calibration studies. Variations of these algorithms with different jet sizes and clustering parameters can easily be configured, as appropriate in the context of a given physics analysis. In addition, other algorithms like the midpoint cone algorithm [28], the seedless infrared safe cone algorithm SIS Cone [6] and all flavours of recursive recombination algorithms provided in the FastJet [8] library, and the “optimal jet finder” described in [29], are available within the standard ATLAS software framework.

Various other jet-related studies are on-going in the ATLAS collaboration in order to understand better and improve jet reconstruction:

- Associating reconstructed tracks with calorimeter signals allows to obtain efficient jet energy corrections. Moreover, vertex information can help in rejecting jets coming from pile-up [24].
- Different studies are on-going in order to understand precisely the effects of pile-up, in particular in the liquid argon calorimeter.
- Jets sub-structure studies such as the use of the “y-scale” given by  $k_t$  jet algorithms [30]

#### 4.4 Conclusion

We gave a brief overview of the recent work of the ATLAS collaboration related to jets reconstruction (details in [31]). In order to deal with the great experimental challenges and to achieve an excellent measurement of hadronic jets, the collaboration has adopted a flexible approach including two main strategies for jet calibration. The collaboration is also preparing several data-based and in-situ techniques to correct and control the jet energy scale and resolution at the precision required by physics analysis as well as conducting several studies to ensure the understanding of the detector response to hadronic jets is optimal.

### 5 b-jets at LHCb

*Author: Victor Coco (on behalf of the LHCb Collaboration)*

LHCb [32] is an LHC experiment dedicated to precise measurements of CP violation and rare B-meson decays. We show that its specifications are of interest for reconstruction and identification of b-jets as well. The LHCb detector is a one-arm spectrometer. It covers the forward region of the interaction point, from 30 mrad to 300 (250) mrad in the bending (non-bending) plane. The choice of such a limited acceptance is motivated by the fact that most of the  $\approx 500 \mu\text{b}$  correlated  $b\bar{b}$  pairs are produced in this region. LHCb experiment will take data at a luminosity of  $2 \times 10^{32} \text{cm}^{-2}\text{s}^{-1}$ , where bunch crossing are dominated by single pp interactions. Good particle identification, excellent tracking and vertexing are needed for B physics measurements. Expected resolution on track momentum is about  $\delta p/p = 0.35\%$  around 10 GeV/c to  $\delta p/p = 0.55\%$  around 140 GeV/c. Impact parameter resolution is expected to be  $\sigma_{\text{IP}} = 14\mu\text{m} + 35\mu\text{m}/p_{\text{T}}$ .

#### 5.1 Reconstruction and identification of b-jets

As a textbook case, we study in the following the case of a Higgs boson decaying into  $b\bar{b}$  pairs, produced in association with a vector boson decaying leptonically. The Higgs mass is chosen

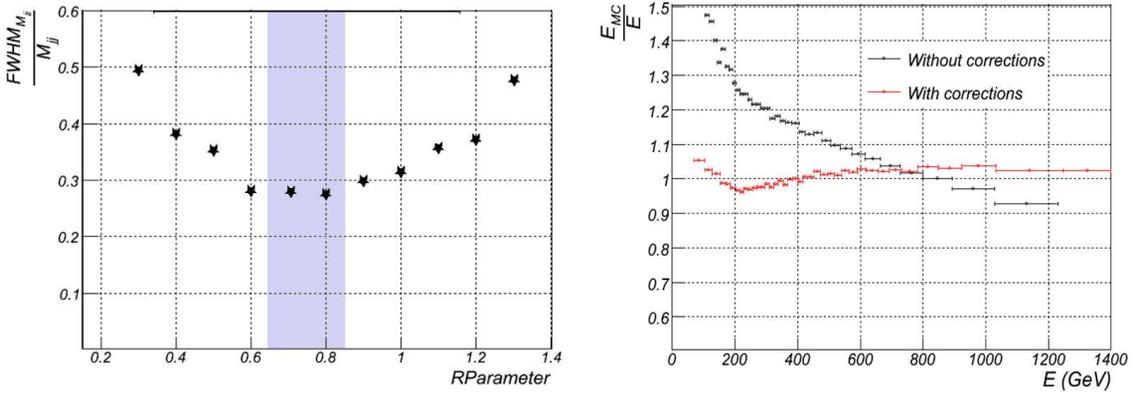


Fig. 15: Full width at half maximum (FWHM) over peak value of the di-jet mass distribution considering all particles from the generator, except neutrinos (left). True jet energy over reconstructed jet energy as a function of the reconstructed jet energy (right).

to be  $120 \text{ GeV}/c^2$  and the lepton, with a transverse momentum  $p_T$  higher than  $10 \text{ GeV}/c$ , is required to be in the LHCb acceptance.

Several contributions might affect the di-jet mass resolution. In order to choose the best working point for the jet algorithm, a generator level study of the di-jet mass reconstruction is performed. The width of the di-jet mass distribution is shown in Fig. 15. At small  $R^2$ , gluon radiation and hadronisation induce a low-mass tail. This effect tends to increase the width of the distribution. For large values of  $R$ , the area of the jets is larger and the contribution of underlying event particles increases. This effect induces high-mass tails, increasing again the width of the distribution. Taking into account both effects, the optimal  $R$  value is between 0.7 and 0.9. The limited acceptance of the detector is the major restriction for jet studies at LHCb. Some jet particles arriving close to the detector border are not reconstructed. Therefore these jets need to be rejected.

Using the LHCb reconstruction framework, charged and neutral particles are separated by matching tracks and electromagnetic calorimeter clusters. In the full simulation, track measurements are used for charged particles while the energy of electromagnetic clusters is used for neutral particles. Tracks and electromagnetic clusters are calibrated objects. In a simple approach, clusters in the hadronic calorimeter are not used. The charged and neutral particles are used as input to the jet algorithm. Two ways of reconstructing and identifying b-jets are under study. The first way consists in finding tracks coming from B-decays, and use them as seed for cone-type jet finding. The second way uses a sequential recombination algorithm,  $k_t$ -jet [1], to reconstruct the jets. the  $k_t$ -jet algorithm with an  $R$  parameter of 0.75 gives on average 15 (proto-)jets for our textbook case events. Only a quarter of them has a transverse momentum higher than  $5 \text{ GeV}/c$ . The performance of b-jets reconstruction is presented below.

A primary cut is applied to remove (uninteresting) jets with  $p_T < 5 \text{ GeV}/c$ , less than 4 constituents and 2% of charged energy. The content of jets in particles from B-decays is

<sup>2</sup> $R$  is a generic parameter of the jet algorithm, representing in first approximation the  $(\eta, \phi)$  radius.

quantified by two weights separating b- from c- and light-jets. To construct these weights, tri-dimensional probability distribution functions have been extracted from Monte Carlo (MC) for particles that are respectively b-, light- and c-jet constituents. b-jet identification is then based on combination of these weights with several variables such as the number of constituents, the energy contained in a cone of  $R < 0.4$  around the jet axis, the charged energy percentage, etc. Taking only into account the jets that pass the primary cut, b-jet selection efficiency is about 81%, for a rejection of about 95% of light-jets and 91% of c-jets in  $t\bar{t}$  events. It is interesting to notice that the b-jets only partially inside the acceptance are rejected at more than 90% by the selection. At this level, no explicit reconstruction of vertices, nor semi-leptonic decay identification has been performed. This leaves room for future improvements.

Reconstruction of b-jets is efficient in the range of pseudorapidity  $2 < \eta < 4.2$ . For jets well contained in the acceptance, one can determine corrections to the jet energy depending on its  $p_T$  and pseudorapidity. The corrections are determined from a  $t\bar{t} \rightarrow b\bar{b} + \ell$  MC sample and applied to the b-jets of the MC Higgs sample. After correction, the energy response presents a non-linearity below 5%, see Fig. 15. In a Gaussian approximation, the di-jet mass resolution is improved from  $\sigma/\text{mean} = 24\%$  to 20%. The mass pick is still offset. The contribution due to loss of neutral hadrons will be added.

The contribution, after b-jets selection, of the remaining "partially inside the acceptance" b-jets, is shown Fig. 16. There is a 10% increase in the resolution of the corrected di-jet mass distribution due to the pollution of those jets.

## 5.2 Interesting processes for LHCb

A measurement of  $H(W, Z) \rightarrow b\bar{b} + \ell$  would be very interesting. But the level of background, especially  $t\bar{t} \rightarrow b\bar{b} + \ell$  is large and its suppression is a real challenge for LHCb. Because of the limited acceptance, global event variables (*e.g.* missing  $E_T$ , sphericity, etc.) are inaccessible. But this analysis also gives the opportunity to develop tools for b-jets studies that might be of interest for other studies involving b-jets in the forward region.

Many new physics models give rise to particles with measurable lifetime decaying into b quarks. In the following we will concentrate on feasibility of two such models.

Hidden valley is a class of phenomenological models that extends the Standard Model (SM) gauge group  $G_{SM}$  with a non-abelian group  $G_V$ . High-dimension operators at the TeV scale allow interactions between SM and new particles. Some neutral v-hadrons,  $\pi_v$ , can decay into the gauge-invariant combinations of SM-particles with observable lifetimes. An interesting process is described by M.J.Strassler and K.M.Zurek [33]. The SM Higgs, thanks to the coupling to a new scalar field, decays into two  $\pi_v$ , which decay into  $b\bar{b}$  pairs. For  $m_{\pi_v} = 45 \text{ GeV}/c^2$ , the probability of correct reconstruction of such a b-jet in LHCb is about 30% up to 50 cm flight path of the  $\pi_v$ . The di-jet mass of b-jets from  $\pi_v$  decay is shown Fig. 17.

Another interesting model was developed by L. Carpenter, D. Kaplan and E-J. Rhee [34]. They have shown that the MSSM with R-parity violation, baryon number violation and non-unified gaugino masses has a non-excluded parameter space in which a light boson decays mainly into two neutralinos. The neutralino decay length varies as the inverse square of the baryon num-

ber violation coupling constant  $\lambda''$ . The final state of such events contains six quarks, among which the probability to find b and c is large. The two vertices from the  $\tilde{\chi}^0 \rightarrow (b, c)$  sequence are reconstructed and assembled to get the  $\tilde{\chi}^0$  vertex mass. A study at generator level with vertex smearing has been performed. For  $m_{\tilde{\chi}^0} = 50 \text{ GeV}/c^2$ ,  $m_{h^0} = 115 \text{ GeV}/c^2$  and  $\lambda'' = 10^{-4}$ , about 16600 events of signal are expected. After selection of the 4 vertices per event, one gets  $\approx 6000$  events, and most of the background is rejected. The  $h^0$  mass distribution after selection, with the remaining events of  $b\bar{b}$ ,  $t\bar{t}$  and  $Z^0 W^\pm \rightarrow b\bar{b}$ , is shown on Fig. 17. Studies of vertex reconstruction and background rejection with full simulation are ongoing. Details can be found in [35].

It has been shown that LHCb can reconstruct b-jets in the forward region ( $2 < \eta < 4$ ), and reconstruct the di-jet mass with a resolution of about 20%. Selection of b-jets benefits a lot from the quality of the LHCb apparatus. Besides important B physics measurements, LHCb has the potential to observe new physics processes in the high rapidity region looking at b-jets and highly displaced vertices.

I would like to thank the LHCb Collaboration and the jets working group for stimulating discussions on the subject and for their help in the preparation of this talk.

## References

- [1] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
- [2] S. D. Ellis and D. E. Soper, Phys. Rev. **D48**, 3160 (1993), arXiv:hep-ph/9305266.
- [3] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997), hep-ph/9707323.
- [4] M. Wobisch and T. Wengler (1998), hep-ph/9907280.
- [5] G. C. Blazey *et al.* (2000), arXiv:hep-ex/0005012.
- [6] G. P. Salam and G. Soyez, JHEP **05**, 086 (2007), arXiv:0704.0292 [hep-ph].
- [7] G. P. Salam and G. Soyez, SISCone, <http://projects.hepforge.org/siscone>.
- [8] M. Cacciari and G. P. Salam, Phys. Lett. **B641**, 57 (2006), arXiv:hep-ph/0512210.
- [9] G. P. S. M. Cacciari and G. Soyez, [www.lpthe.jussieu.fr/~salam/fastjet](http://www.lpthe.jussieu.fr/~salam/fastjet).
- [10] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008), arXiv:0802.1189 [hep-ph].
- [11] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006), arXiv:hep-ph/0603175.
- [12] C. Buttar *et al.* (2008), arXiv:0803.0678 [hep-ph].

- [13] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, *Phys. Lett.* **B269**, 432 (1991).
- [14] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, *Phys. Rev. Lett.* **100**, 242001 (2008), [arXiv:0802.2470 \[hep-ph\]](#).
- [15] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, *JHEP* **12**, 032 (2008), [arXiv:0810.1304 \[hep-ph\]](#).
- [16] M. Dasgupta, L. Magnea, and G. P. Salam, *JHEP* **02**, 055 (2008), [arXiv:0712.3014 \[hep-ph\]](#).
- [17] A. Bhatti *et al.* (2008), [arXiv:0807.4961 \[hep-ex\]](#).
- [18] M. Cacciari and G. P. Salam, *Phys. Lett.* **B659**, 119 (2008), [arXiv:0707.1378 \[hep-ph\]](#).
- [19] M. Cacciari, G. P. Salam, and G. Soyez, *JHEP* **04**, 005 (2008), [arXiv:0802.1188 \[hep-ph\]](#).
- [20] T. C. Collaboration, CMS PAS JME-07-003 (2008).
- [21] ATLAS Collaboration, S. Bentvelsen *et al.*, *JINST* **3**, S08003 (2008).
- [22] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics – Jets and Missing Et Chapter, Jet Reconstruction Performance*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [23] W. e. a. Lampl, *Calorimeter Clustering Algorithms: Description and Performance*, Technical Report ATL-LARG-PUB-2008-002, CERN, Geneva, Apr 2008.
- [24] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics – Jets and Missing Et Chapter, Detector Level Jet Corrections*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [25] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics – Jets and Missing Et Chapter, E/p Performance for Charged Hadrons*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [26] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics – Jets and Missing Et Chapter, In-situ Calibration Strategies*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [27] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics – Top-quark Physics at ATLAS CSC Chapter*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [28] A. P. Cheplakov and S. Thompson, *MidPoint Algorithm for Jets Reconstruction in ATLAS Experiment*, Technical Report ATL-PHYS-PUB-2007-007, CERN, Geneva, Dec 2006.

- [29] D. Y. Grigoriev, E. Jankowski, and F. V. Tkachov, *Comput. Phys. Commun.* **155**, 42 (2003), [arXiv:hep-ph/0301226](#).
- [30] J. M. Butterworth, A. Davison, E. Ozcan, and P. Sherwood, *YSplitter: An Athena tool for studying jet substructure*, Technical Report ATL-PHYS-INT-2007-015. ATL-COM-PHYS-2007-077, CERN, Geneva, Oct 2007.
- [31] A. Collaboration, *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*, No. CERN-OPEN-2008-020. CERN, 2008, to appear.
- [32] LHCb Collaboration, S. Amato *et al.* CERN-LHCC-98-04.
- [33] M. J. Strassler and K. M. Zurek, *Phys. Lett.* **B651**, 374 (2007), [arXiv:hep-ph/0604261](#);  
M. J. Strassler and K. M. Zurek, *Phys. Lett.* **B661**, 263 (2008), [arXiv:hep-ph/0605193](#).
- [34] D. E. Kaplan and K. Rehermann, *JHEP* **10**, 056 (2007), [arXiv:0705.3426](#) [hep-ph];  
L. M. Carpenter, D. E. Kaplan, and E.-J. Rhee, *Phys. Rev. Lett.* **99**, 211801 (2007), [arXiv:hep-ph/0607204](#).
- [35] N. Gueissaz. CERN-THESIS-2007-038.

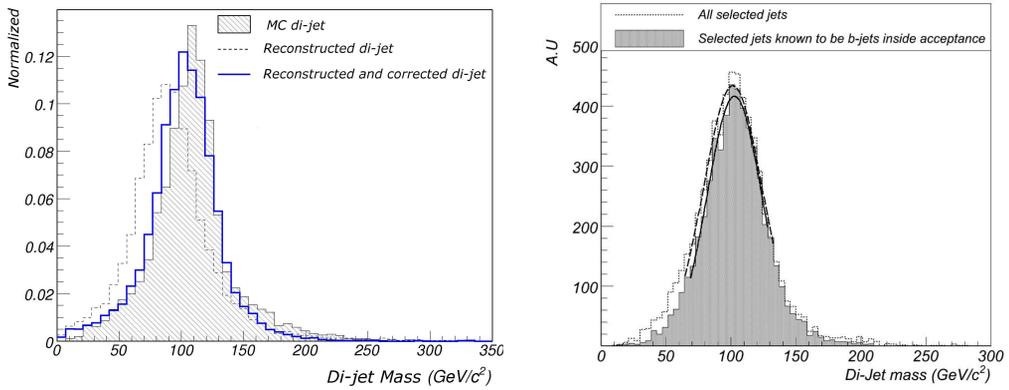


Fig. 16: Full simulation di-jet mass with and without correction, (left). Corrected di-jet mass with "inside the acceptance" b-jets only (filled grey), and all selected b-jets (dashed line), (right).

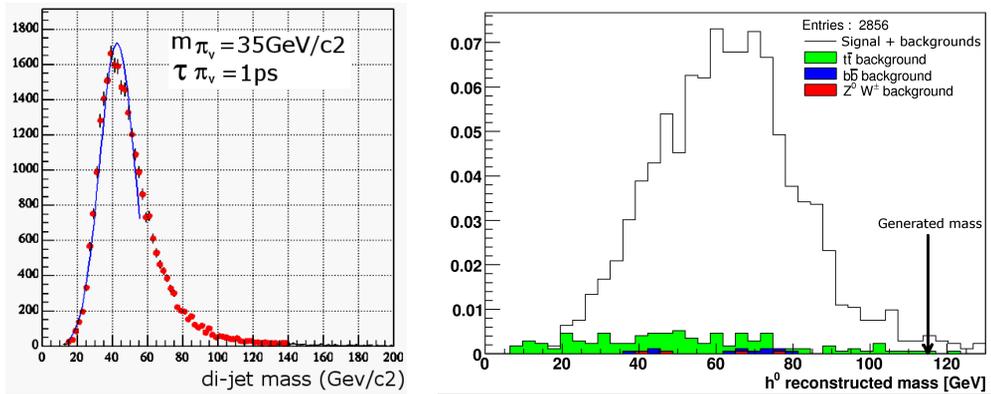


Fig. 17: Reconstructed  $\pi_\nu$  mass from Higgs decays with  $m_H = 120 \text{ GeV}/c^2$  in the hidden valley model (left). Higgs mass reconstructed from the two  $\tilde{\chi}^0$  vertices in MSSM with an R-parity violation model (right).

# $k_{\perp}$ -factorization and forward jets

*Sergey Baranov, Jochen Bartels, Michal Deák, Francesco Hautmann, Hannes Jung, Albert Knutsson, Krzysztof Kutak, Artem Lipatov, Christophe Royon, Augustín Sabio Vera, Florian Schwennsen, Nikolai Zotov*

Hadronic final states containing multiple jets have been investigated at the Tevatron and HERA colliders, and will play a central role in the Large Hadron Collider (LHC) physics program. The interpretation of experimental data for such final states relies both on perturbative multi-jet calculations (see [1] for a recent overview) and on realistic event simulation by parton-shower Monte Carlo generators (see e.g. [2–6]).

Owing to the complex kinematics involving multiple hard scales and the large phase space opening up at very high energies, multi-jet events are potentially sensitive to effects of QCD initial-state radiation that depend on the finite transverse-momentum tail of partonic matrix elements and distributions.

Standard shower Monte Carlos reconstructing exclusive events, such as HERWIG [7,8] and PYTHIA [9], are based on collinear evolution of the initial-state jet. Finite- $k_{\perp}$  contributions are not included, but rather correspond to corrections [10–14] to the angular or transverse-momentum ordering implemented in the parton-branching algorithms. The theoretical framework to take these corrections into account is based on using initial-state distributions unintegrated in both longitudinal and transverse momenta [12–14], coupled to hard matrix elements (ME) suitably defined off mass shell. See e.g. [15] for discussion of the Monte Carlo shower implementation of the method. Event generators based on  $k_{\perp}$ -dependent showers of this kind include [16–22].

We give a short introduction to  $k_{\perp}$ -factorization and describe the determination of unintegrated parton density functions (uPDFs). Then we discuss the calculation of new processes in the frame of  $k_{\perp}$ -factorisation and show a comparison with measurements at the Tevatron. A summary of NLO calculations for multi-jet production in  $ep$  and  $pp$  in  $k_{\perp}$ -factorisation follows. Finally, we discuss forward jet production and the azimuthal decorrelation of jets both in  $ep$  and  $pp$ , signatures which could clearly show evidence for small  $x$  parton dynamics.

## 1 Short introduction to $k_{\perp}$ -factorization and uPDFs

*Author: Francesco Hautmann, Hannes Jung*

In  $k_{\perp}$ -factorization the cross section for any process  $pp \rightarrow X$  can be written as:

$$\sigma = \int dx_1 dx_2 \int dk_{\perp 1} dk_{\perp 2} \mathcal{A}(x_1, k_{\perp 1}, q) \mathcal{A}(x_2, k_{\perp 2}, q) \hat{\sigma}(x_1, x_2, k_{\perp 1}, k_{\perp 2}, q) \quad (1)$$

with  $\mathcal{A}(x, k_{\perp}, q)$  being the un-integrated transverse momentum ( $k_{\perp}$ -dependent) parton density function (uPDF or TMD),  $q$  defines the factorization scale and  $\hat{\sigma}$  is the partonic cross section taken with off-shell initial partons. Here we concentrate on the small  $x$  region, which is described by high-energy factorization (or  $k_{\perp}$ -factorisation).

Both the uPDF and the off-shell partonic cross section can be formulated in the small  $x$  region where a gauge-invariant definition emerges from high-energy factorization [12–14]. It has

been used for studies of collider processes both by Monte Carlo (see reviews in [23–25]) and by semi-analytic resummation approaches (see [26, 27]).

To characterize a transverse momentum dependent parton distribution gauge-invariantly over the whole phase space is a nontrivial question [28, 29], currently at the center of much activity. See overview in [24].

The diagrammatic argument for gauge invariance, given in [12–14], and developed in [30, 31], is based on relating off-shell matrix elements with physical cross sections at  $x \ll 1$ , and exploits the dominance of single gluon polarization at high energies.<sup>1</sup> The main reason why a natural definition for TMD pdfs can be constructed in the high-energy limit is that one can relate directly (up to perturbative corrections) the cross section for a *physical* process, say, photoproduction of a heavy-quark pair, to an *unintegrated* gluon distribution, much as, in the conventional parton picture, one does for DIS in terms of ordinary (integrated) parton distributions. On the other hand, the difficulties in defining a TMD distribution over the whole phase space can largely be associated with the fact that it is not obvious how to determine such a relation for general kinematics.

The evolution equations obeyed by TMD distributions defined from the high-energy limit are of the type of energy evolution [32]. Factorization formulas in terms of TMD distributions [12–14] have corrections that are down by logarithms of energy rather than powers of momentum transfer. On the other hand, it is important to observe that this framework allows one to describe the ultraviolet region of arbitrarily high  $k_{\perp}$  and in particular re-obtain the structure of QCD logarithmic scaling violations [26, 27, 30, 31]. This ultimately justifies the use of this approach for jet physics. In particular it is the basis for using corresponding Monte Carlo implementations [15–22] to treat multi-scale hard processes at the LHC.

From both theoretical and phenomenological view-points, it is one of the appealing features of the high-energy framework for TMD distributions that one can relate its results to a well-defined summation of higher-order radiative corrections. By expanding these results to fixed order in  $\alpha_s$ , one can match the predictions thus obtained against perturbative calculations. This has been verified for a number of specific processes at next-to-leading order (see for instance [33–35] for heavy flavor production) and more recently at next-to-next-to-leading order (see for instance [36, 37]). Note that this fact also provides the basis for shower algorithms implementing this framework to be combined with fixed-order NLO calculations by using existing techniques for such matching.

## 2 Prospects and recent developments of $k_{\perp}$ -factorization

At HERA the  $k_{\perp}$ -factorization approach has been successfully applied to describe multi-jet production as well as the production of heavy quarks at small values of  $x$ , which are dominated by gluon initiated processes. The relevant off-shell matrix elements for jet and heavy quark production are known since long. The unintegrated gluon distribution has been determined using inclusive measurements at HERA. A new determination of the uPDF using also final state measurements is described in section 2.1.

---

<sup>1</sup>It is emphasized e.g. in [23, 29] that a fully worked out operator argument, on the other hand, is highly desirable but is still missing.

However to apply  $k_{\perp}$ -factorization to describe measurements in general in  $p\bar{p}$  or  $pp$  new and additional matrix elements for different processes need to be calculated. In the following, the calculation of new processes will be presented:

- $g^*q \rightarrow gq$  to describe jet production in the forward and backward region
- $g^*g^* \rightarrow \gamma/W/Z + q\bar{q}$  to describe the inclusive production of  $\gamma/W/Z$
- $g^*q \rightarrow \gamma q$  to describe prompt photon production

Since some of the processes are quark initiated, unintegrated quark densities need to be determined. In a simplest approach we allow only valence quarks (at large  $x$ ). The contribution of quark initiated processes is discussed in section 2.3 explicitly.

The aim of this contribution is to show the two areas, where improvements in the  $k_{\perp}$ -factorization approach has been made: the determination of the uPDFs and the calculation of matrix elements.

## 2.1 An approach to fast fits of the unintegrated gluon density

*Author: Alessandro Bacchetta, Albert Knutsson, Krzysztof Kutak*

In perturbative QCD the PDFs are given by solutions of integral equations, for which the initial input distributions have to be determined by fits to experimental data. It turns out that, in general it is not efficient to tune Monte Carlo event generators (MC) by sequential calls of the generator together with a minimisation program. Motivated by [38], we use an alternative fitting method, which is based on producing a grid in parameter-observable space. This allows the parameter dependence to be determined by polynomial interpolation before the fit is performed, which significantly reduces the time to do the fit itself.

Here we determine the parameters in the starting distribution of the unintegrated gluon density function by fits to deep inelastic scattering  $F_2$  structure function data from the H1 experiment [39]. This is carried out by using the CASCADE Monte Carlo event generator [16].

### *The Unintegrated Gluon Density*

The starting distribution of the unintegrated gluon density is parameterized as

$$A_0(x, k_t) = Nx^{-B}(1-x)^C(1-Dx) \exp [(k_t - \mu)^2/\sigma^2]$$

where  $x$  is the longitudinal momentum fraction of the proton carried by the gluon and  $k_t$  its transverse momentum. In this study the  $N$  (normalisation),  $B$  (low  $x$  behaviour),  $D$  are determined. The parameters  $C$ ,  $\sigma$  and  $\mu$ , are kept fixed at  $C = 4$ ,  $\sigma = 1$  and  $\mu = 0$ .

The unintegrated gluon density is determined by a convolution of the non-perturbative starting distribution  $\mathcal{A}_0(x)$  and the CCFM evolution denoted by  $\tilde{\mathcal{A}}(x, k_{\perp}, \bar{q})$ :

$$x\mathcal{A}(x, k_{\perp}, \bar{q}) = \int dx' \mathcal{A}_0(x', k_{\perp}) \cdot \frac{x}{x'} \tilde{\mathcal{A}}\left(\frac{x}{x'}, k_{\perp}, \bar{q}\right)$$

### The Fitting Method

In the first step of the fitting procedure we build up a grid of MC predictions in the parameter space  $(p_1, p_2, \dots, p_n)$  for each of the observables  $X$ . Then we use the grid to describe the parameter space analytically by a polynomial of the form

$$X(p_1, p_2, \dots, p_n) = A_0 + \sum_{i=1}^n B_i p_i + \sum_{i=1}^n C_i p_i^2 + \sum_{i=1}^{n-1} \sum_{j=i+1}^n D_{ij} p_i p_j + \text{H.O.}$$

We determine the coefficients  $A_0, B_1, \dots$  by singular value decomposition (SVD) [40], since they form an over determined system of linear equations. This is done separately for each of the MC predicted observables, which in our case corresponds to 58 experimental data points.

In order to account for correlations between parameters the form of the polynomial has to be of order higher than one. In the presented fit we use a fourth order polynomial, which gives a good description of the parameter space. The  $\chi^2/n.d.f.$ , averaged over the 58 MC predicted observables, is  $\chi^2/n.d.f. = 501.7/(440 - 35) = 1.2$ , where 440 is the number of MC grid points and 35 is the number of coefficients in the polynomial of the fourth degree.

Having described with the polynomials the behavior of the MC predictions in parameter space, we can find the values of the parameters  $p_1, p_2, \dots$  for which the MC best reproduces the measurements. This is done by applying a  $\chi^2$  minimisation to

$$\chi^2 = \sum_k \frac{(X_{k,poly} - X_{k,data})^2}{\delta X_{k,poly}^2 + \delta X_{k,data}^2}$$

where the sum runs over all bins,  $k$ .  $X_{k,data}$  is the measured data, with the corresponding experimental error  $\delta X_{k,data}$ , and  $X_{k,poly}$  the polynomial prediction, with the error  $\delta X_{k,poly}$  calculated from the individual errors of the fitted coefficients by using the covariance matrix. To perform this last step we use MINUIT [41], since the dependence on the parameters  $p_1, p_2, \dots$  is non-linear.

The method turns out to be very time-efficient, in particular since the MC grid points are generated simultaneously.

### Results

The unintegrated PDF has been fitted to the proton structure function,  $F_2$ , in the kinematical range  $Q^2 > 4.5 \text{ GeV}^2$  and  $x_{Bj} < 0.005$ , where  $Q^2$  is the virtuality of the exchanged boson and  $x_{Bj}$  is the Bjorken scaling variable. In Fig. 1 the result of the fit is compared to the data [39]. The parameter values determined from the new fit are  $N = 0.221 \pm 0.011$ ,  $B = 0.201 \pm 0.007$  and  $D = -24.6 \pm 1.5$ .  $\chi^2$  profiles for these parameters are shown in Fig. 2 and confirm that this is a minimum for  $N, B$  and  $D$ . Confidence regions for these parameters are shown in Fig. 3.

The  $\chi^2/ndf$  of the new fit is 2.4 which is more than one unit better than obtained when using CASCADE together with the PDF set A0 [16]. The constraints on the parameters of the gluon could be hopefully further improved by fitting the  $k_t$ -dependent part of the gluon distribution. In Fig. 4 the new gluon distribution and set A0 are drawn as a function of  $x$  for two different values of  $k_t^2$ .

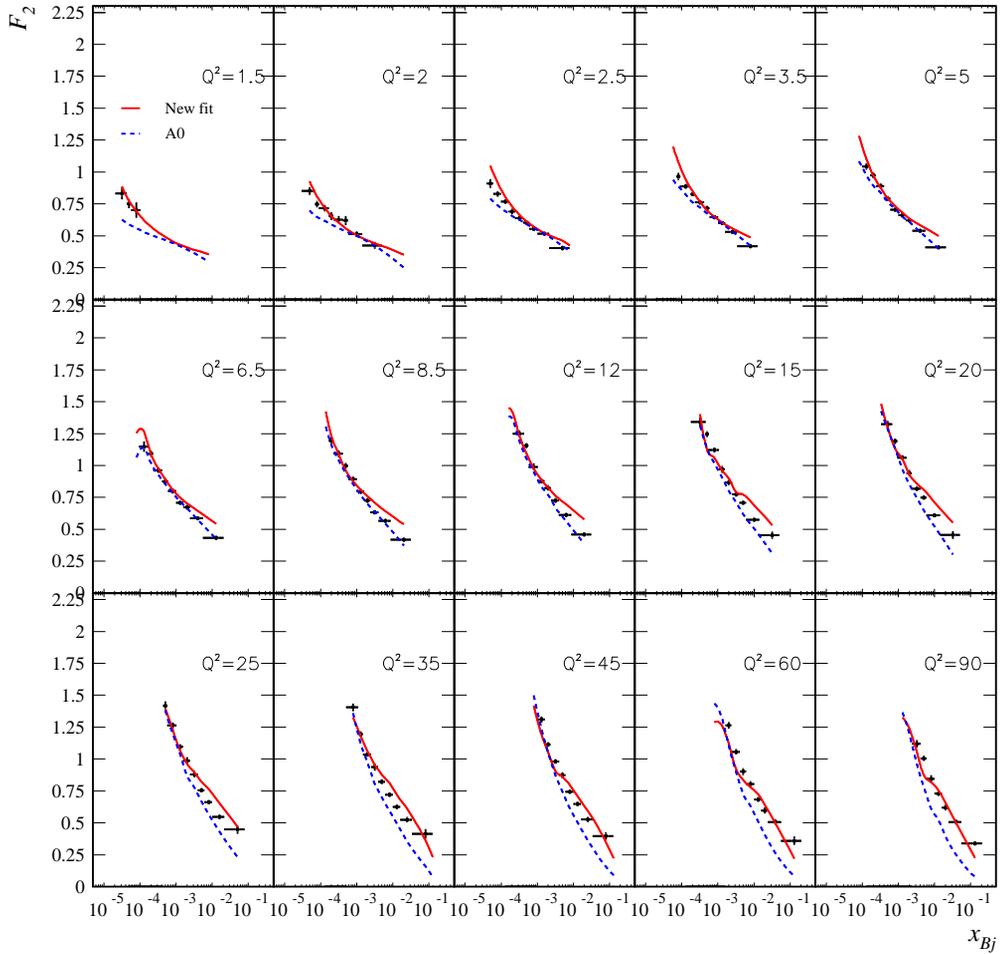
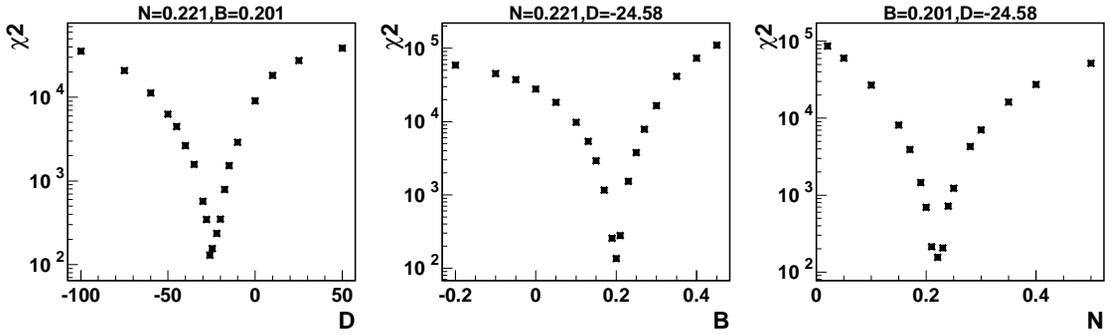
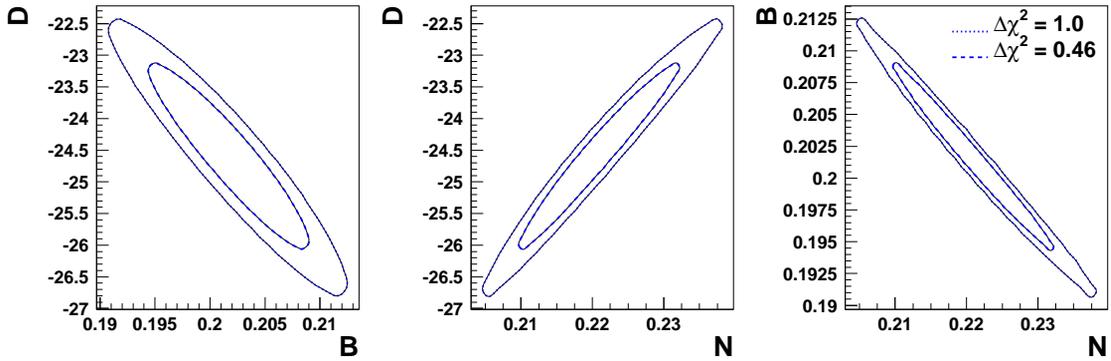


Fig. 1:  $F_2$  structure function data, as a function of  $x_{Bj}$  in  $Q^2$  bins compared to predictions from the CASCADE Monte Carlo event generator using the newly fitted PDF (continuous red line) and old PDF set A0 (blue dashed line). The new fitted PDF has been determined in the kinematic range  $Q^2 > 4.5 \text{ GeV}^2$  and  $x_{Bj} < 0.005$

Fig. 2:  $\chi^2$  as a function of the fitted parameters.Fig. 3: Confidence regions for  $N$ ,  $B$  and  $D$ 

## 2.2 Hard matrix element calculation

For jet production at high energies the following processes contribute:  $gg \rightarrow q\bar{q}$ ,  $gg \rightarrow gg$ ,  $qg \rightarrow qg$ ,  $qq \rightarrow qq$  and  $q\bar{q} \rightarrow q\bar{q}$ . In collinear factorization (with on-shell initial partons) these processes are calculated in LO ( $\mathcal{O}(\alpha_s)$ ) and also higher order corrections are known. In  $k_\perp$ -factorization the process  $gg \rightarrow q\bar{q}$  are known [13, 42]. At high energies, gluon induced processes are expected to dominate. The process  $g^*g^* \rightarrow gg$  is not yet considered, as there will be contributions of similar type from the parton branching. However, if jet production in the forward or backward region is considered, scattering a small  $x$  gluon off a large  $x$  valence quark ( $qg^* \rightarrow qg$ ) will contribute significantly. This process will be described below.

The production of  $Z/W$  is calculated to a high precision in collinear factorization, even to NNLO. However, significant effects from small  $x$  partons, which are not included in the collinear treatment could become important, as suggested by [43]. Since  $W/Z$  production is the standard candle at LHC, it is important to understand in detail any possible small  $x$  effect. The  $Z/W$  production has been calculated for the first time in the framework of  $k_\perp$ -factorization in [44, 45] for the lowest order gluon induced process  $g^*g^* \rightarrow Z/W + Q_i\bar{Q}_j$ . In [45] attempts are made to include also quark initiated processes to  $Z/W$  production.

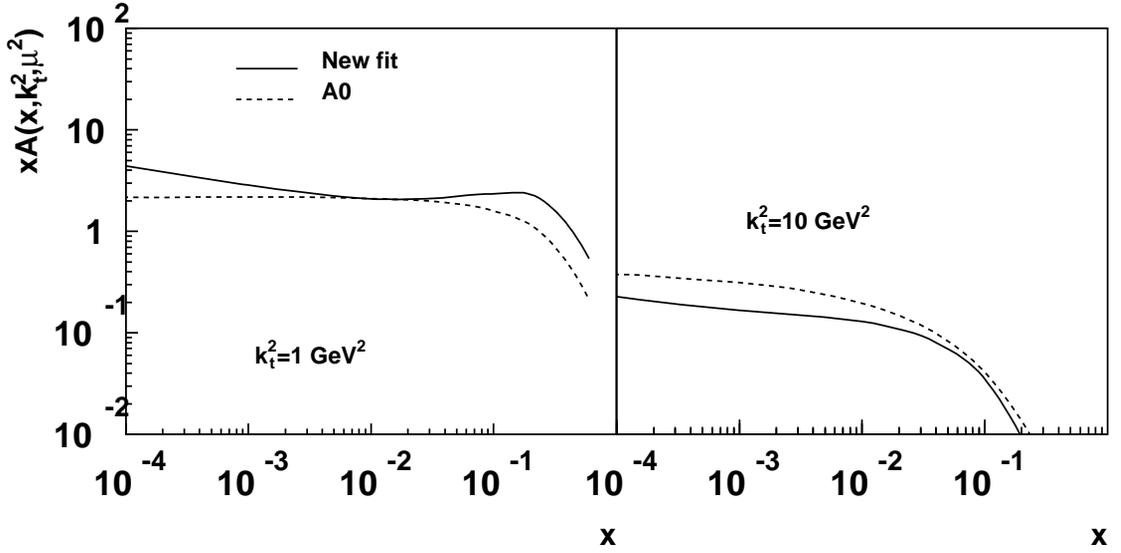


Fig. 4: The newly fitted PDF (continuous line) compared to the old PDF set A0 (dashed line), drawn as a function of  $x$  for  $k_t^2 = 1 \text{ GeV}^2$  and  $k_t^2 = 10 \text{ GeV}^2$  at a scale  $\mu^2 = 4 \text{ GeV}^2$ .

### Process $qg^* \rightarrow qg$ at the LHC

Author: Michal Deak, Krzysztof Kutak

Here we consider a very asymmetric situation in proton proton scattering in which an off-shell gluon coming from one of the protons scatters off an on-shell valence quark from the other proton. We can use small  $x$  dynamics for the gluon where the  $k_{\perp}$ -factorization formalism is justified and on the other side we use collinear large  $x$  dynamics for the valence quark. The matrix element of the hard subprocess is factorized from the unintegrated gluon density function by  $k_{\perp}$ -factorization theorem and from the valence quark uPDF.

Similar to the  $Z/W + Q_i \bar{Q}_j$  case ([44] and [45]), we will use Sudakov decomposition for the four-momenta of the initial state and final state particles.

$$k = x_g p_A + z_g p_B + k_{\perp} \quad (2)$$

$$q = x_q p_B \quad (3)$$

$$k' = x'_g p_A + z'_g p_B + k'_{\perp} \quad (4)$$

$$q' = z'_q p_A + x'_q p_B + q'_{\perp} \quad (5)$$

$$t = (k - k')^2 \quad (6)$$

The amplitude for the process  $g^* q \rightarrow gq$  consists of the diagrams in Fig. 5. The squared matrix element, after summing over colors of final and initial state particles, is calculated using the  $k_{\perp}$ -factorization prescription:

$$|\mathcal{M}|^2 = \frac{1}{4 N_c (N_c^2 - 1)} \left( C_A C_F^2 A_{abelian} + C_A^2 C_F A_{nonabelian} \right) \quad (7)$$

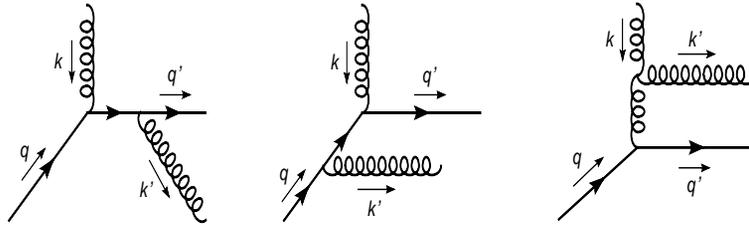


Fig. 5: Full set of diagrams of  $qg^* \rightarrow qg$  with initial state gluon off-shell required by gauge invariance. a) Diagrams similar to diagrams in collinear factorization approach, b) additional diagrams required by gauge invariance.

where

$$A_{abelian} = (4\pi\alpha_s)^2 \left( \frac{k \cdot q}{p \cdot q} \right)^2 \frac{(p \cdot q)^2 + (p \cdot q')^2}{k' \cdot q k' \cdot q'} \quad (8)$$

and

$$A_{nonabelian} = (4\pi\alpha_s)^2 \left( \frac{k \cdot q}{p \cdot q} \right)^2 \frac{(p \cdot q)^2 + (p \cdot q')^2}{2k' \cdot q k' \cdot q'} \left( \frac{2k' \cdot q' p \cdot q}{-t k' \cdot p} + \frac{2k' \cdot q p \cdot q'}{-t k' \cdot p} - 1 \right) \quad (9)$$

with  $C_A = N_c$ ,  $C_F = (N_c^2 - 1)/(2N_c)$  and  $N_c$  being the number of colours. The  $k_\perp \rightarrow 0$  can be performed and the text book result for  $qg \rightarrow qg$  is recovered.

The matrix element is singular when one of the particles in final state is collinear with the quark in initial state. To regularize the matrix element we set a cut on the transverse momenta of each of final state particles in the laboratory frame,  $|\mathbf{k}'| > p_{\perp cut}$  and  $|\mathbf{q}'| > p_{\perp cut}$ . We note that a cut on one of the transverse momenta is not enough to avoid divergencies.

### *Z and W production associated with heavy quark-antiquark pair*

The calculation of the matrix element for the process  $g^*g^* \rightarrow Z/W + Q_i\bar{Q}_j$  is described in detail in [44,45]. The calculations differ in the way the spin density of the initial state is treated. However, they are equivalent and give the same results for the matrix element of the hard subprocess. We have cross-checked the calculations numerically and found agreement of the cross sections at Tevatron and LHC energies at the 0.1 % level.

## 2.3 Implications for the LHC: Electroweak gauge boson production in hadronic collisions at high energies

*Author: Serguei Baranov, Artem Lipatov, Nikolai Zotov*

AT HERA and the Tevatron  $k_\perp$ -factorization supplemented with the BFKL-like gluon dynamics was successfully applied to describe various measurements of heavy quark production [46, 47] (and references therein). It is important that these predictions were based on the off-shell matrix elements  $\gamma g^* \rightarrow Q\bar{Q}$  or  $g^*g^* \rightarrow Q\bar{Q}$ . In Ref. [48, 49] inclusive Higgs hadroproduction at Tevatron and LHC energies has been investigated, where the main contribution also

came from the off-shell gluon-gluon fusion. It was demonstrated that using the CCFM-evolved unintegrated gluon densities results in predictions which are very close to the NNLO pQCD ones. This encouraged us to apply the  $k_\perp$ -factorization approach also to the production of inclusive electroweak gauge bosons.

At leading order (LO) QCD, the  $W^\pm$  and  $Z^0$  bosons are produced via quark-antiquark annihilation  $q + \bar{q}' \rightarrow W/Z$ . Here, an important component of the calculations are the unintegrated quark distributions. At present, these distributions are only available in the Kimber-Martin-Ryskin (KMR) scheme [50, 51], since there are theoretical difficulties in obtaining quark distributions directly from CCFM and BFKL equations. This is in contrast to gluon-induced processes where many unintegrated gluon densities are available.

Since sea quarks can appear as a result of gluon splitting, at the price of absorbing the last gluon splitting into the hard subprocess (i.e., considering the  $2 \rightarrow 2$  and  $2 \rightarrow 3$  rather than  $2 \rightarrow 1$  matrix elements), the problem of poorly known sea quark densities can efficiently be reduced to the problem of gluon densities. However, it is not evident in advance whether the last gluon splitting dominates. This issue is addressed in Ref. [45, 52]. One of the goals of that study is to clarify, to what extent the quark contributions can be reexpressed in terms of the gluon contributions. At the same time, by considering the higher order matrix elements we take into account the terms not containing large logarithms, i.e., the terms not included in the evolution equations. Within our scheme, we get a numerical estimate of the corresponding contributions.

Our theoretical approach is the following. We start from the leading order  $\mathcal{O}(\alpha)$  subprocess  $q + \bar{q}' \rightarrow W/Z$ , and then divide it into several contributions which correspond to the interactions of valence quarks  $q_v(x, \mathbf{k}_T^2, \mu^2)$ , sea quarks appearing at the last step of the gluon evolution  $q_g(x, \mathbf{k}_T^2, \mu^2)$ , and sea quarks coming from the earlier steps  $q_s(x, \mathbf{k}_T^2, \mu^2)$ . Here we use the specific property of the KMR scheme which enables us to discriminate between the various components of the quark densities.

The KMR approach represents an approximate treatment of the parton evolution mainly based on the DGLAP equation and incorporating BFKL effects at the last step of the parton ladder only, in the form of properly defined Sudakov formfactors  $T_q(\mathbf{k}_T^2, \mu^2)$  and  $T_g(\mathbf{k}_T^2, \mu^2)$ . These formfactors already include logarithmic loop correction. Also, there are nonlogarithmic corrections which result in a K-factor on the cross section given by [53]  $K(q + \bar{q}' \rightarrow W/Z) \simeq \exp [C_F \pi \alpha_s(\mu^2)/2]$  with  $C_F = 4/3$  and  $\mu^2 = \mathbf{p}_T^{4/3} m^{2/3}$ . In this approximation, the unintegrated quark and gluon distributions are expressed by

$$f_q(x, \mathbf{k}_T^2, \mu^2) = T_q(\mathbf{k}_T^2, \mu^2) \frac{\alpha_s(\mathbf{k}_T^2)}{2\pi} \times \int_x^1 dz \left[ P_{qq}(z) \frac{x}{z} q\left(\frac{x}{z}, \mathbf{k}_T^2\right) \Theta(\Delta - z) + P_{qg}(z) \frac{x}{z} g\left(\frac{x}{z}, \mathbf{k}_T^2\right) \right], \quad (10)$$

$$f_g(x, \mathbf{k}_T^2, \mu^2) = T_g(\mathbf{k}_T^2, \mu^2) \frac{\alpha_s(\mathbf{k}_T^2)}{2\pi} \times \int_x^1 dz \left[ \sum_q P_{gq}(z) \frac{x}{z} q\left(\frac{x}{z}, \mathbf{k}_T^2\right) + P_{gg}(z) \frac{x}{z} g\left(\frac{x}{z}, \mathbf{k}_T^2\right) \Theta(\Delta - z) \right], \quad (11)$$

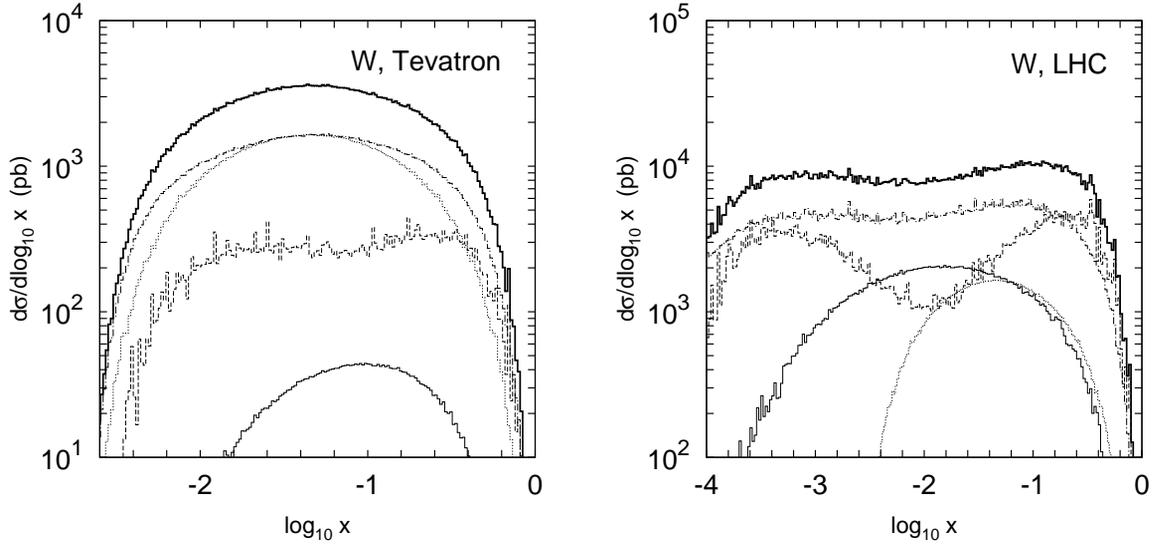


Fig. 6: Different contributions to the inclusive  $W^\pm$  boson production at the Tevatron (left panel) and LHC (right panel) conditions. The solid, dashed and dotted histograms represent the contributions from the  $g^* + g^* \rightarrow W^\pm/Z^0 + q + \bar{q}'$ ,  $q_v + g^* \rightarrow W^\pm/Z^0 + q'$  and  $q_v + \bar{q}'_v \rightarrow W^\pm/Z^0$  subprocesses, respectively. The dash-dotted histograms represent the “reduced sea” component. The thick solid histograms represent the sum of all contributions.

where  $P_{ab}(z)$  are the usual unregularised leading order DGLAP splitting functions, and  $q(x, \mu^2)$  and  $g(x, \mu^2)$  are the conventional (collinear) quark and gluon densities. The function  $f_q(x, \mathbf{k}_T^2, \mu^2)$  in Eq. (10) represents the total quark distribution. Modifying Eq. (10) in such a way that only the first term is kept and the second term omitted, we switch the last gluon splitting off, thus excluding the  $q_g(x, \mathbf{k}_T^2, \mu^2)$  component. Taking the difference between the quark and antiquark densities we extract the valence quark component  $q_v(x, \mathbf{k}_T^2, \mu^2) = f_q(x, \mathbf{k}_T^2, \mu^2) - f_{\bar{q}}(x, \mathbf{k}_T^2, \mu^2)$ .

Summing up, we consider the following partonic subprocesses: gluon-gluon fusion  $g + g \rightarrow W/Z + q + \bar{q}'$ , with which the  $q_g + \bar{q}_g$  annihilation is replaced; valence and sea quark-gluon scattering  $q_v + g \rightarrow W/Z + q$  and  $q_s + g \rightarrow W/Z + q$ , with which the  $q_v + \bar{q}_g$  and  $q_s + \bar{q}_g$  annihilation is replaced; and quark-antiquark annihilation  $q + \bar{q}' \rightarrow W/Z$  including both valence  $q_v$  and sea  $q_s$  quark components. The calculation of the matrix elements is explained in section 2.2. The basic formulas for corresponding contributions to the cross section are given in [45, 52].

Now we turn to numerical results. The solid, dashed and dotted histograms in fig. 2.3 represent the contributions from the  $g^* + g^* \rightarrow \gamma/W^\pm/Z^0 + q + \bar{q}'$ ,  $q_v + g^* \rightarrow \gamma/W^\pm/Z^0 + q'$  and  $q_v + \bar{q}'_v \rightarrow W^\pm/Z^0$  (or  $q_v + \bar{q}_v \rightarrow \gamma + g$ ) subprocesses, respectively. The dash-dotted histograms represent the sum of the contributions from the  $q_s + \bar{q}'_s \rightarrow W^\pm/Z^0$ ,  $q_s + g^* \rightarrow \gamma/W^\pm/Z^0 + q'$  and  $q_v + \bar{q}'_s \rightarrow W^\pm/Z^0$  (or  $q_s + \bar{q}_s \rightarrow \gamma + g$  and  $q_v + \bar{q}_s \rightarrow \gamma + g$ ) subprocesses. We find that the contribution from the valence quark-antiquark annihilation is important at the Tevatron but yields only about few percent at the LHC energy. The gluon-gluon fusion is unimportant at the

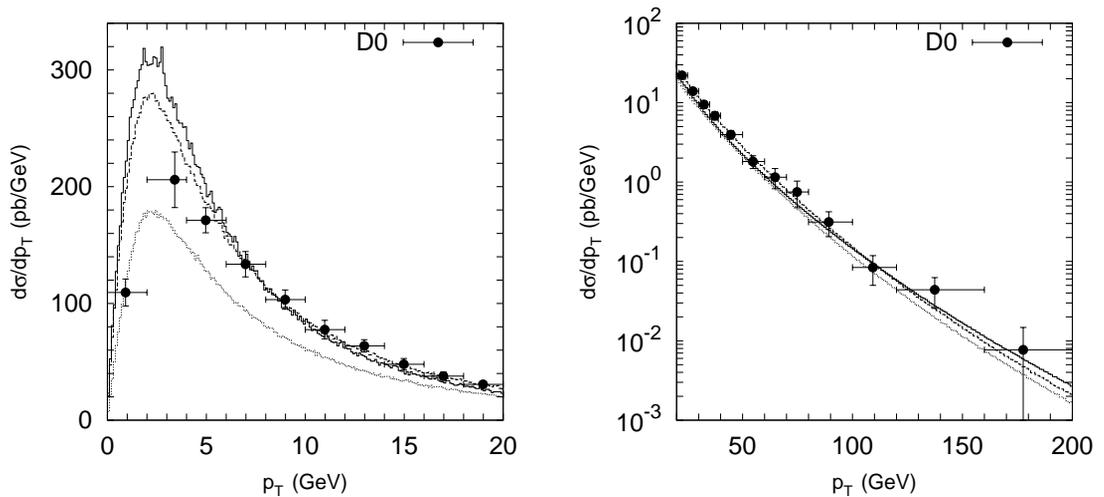


Fig. 7: Transverse momentum distribution of the  $W^{\pm}$  boson production calculated at  $\sqrt{s} = 1800$  GeV. Solid histograms represent calculations in the "decomposition" scheme where all contributions described in the text are taken into account. Dashed histograms correspond to the predictions based on the simple  $2 \rightarrow 1$  quark-antiquark annihilation subprocess with all quark components summed together. Dotted histograms correspond to the simple  $2 \rightarrow 1$  quark-antiquark annihilation subprocess without  $K$ -factor. The cross sections times branching fraction  $f(W \rightarrow l\nu)$  are shown. The experimental data are from D0.

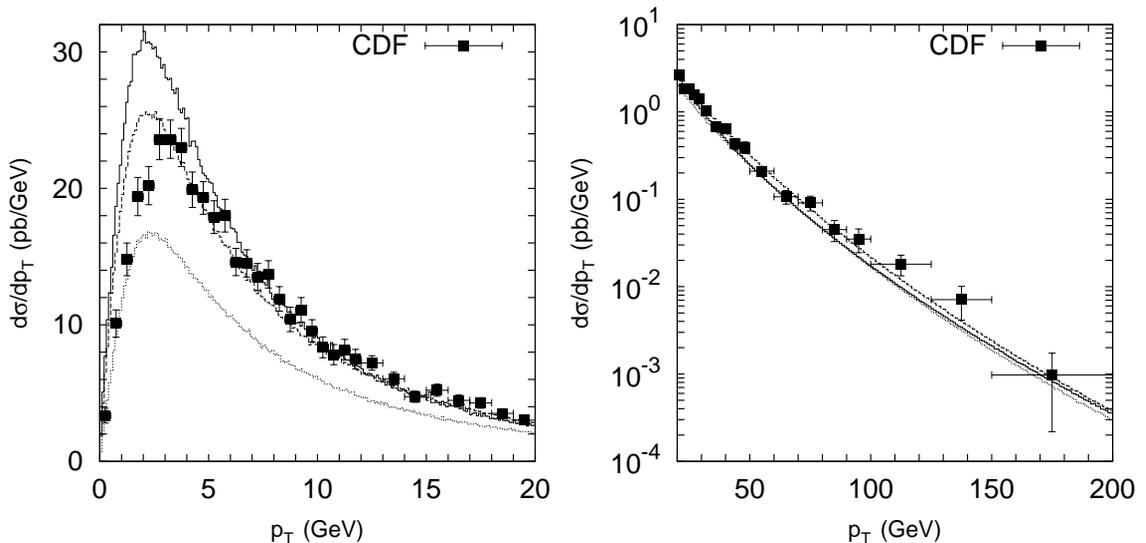


Fig. 8: Transverse momentum distribution of the  $Z^0$  boson. Notation of the histograms is as in Fig.2.3. The experimental data are from CDF.

Tevatron, but becomes important at higher energies and has to be taken into account at the LHC. Quite a significant fraction (nearly 50%) of the calculated cross section at both the Tevatron and the LHC conditions comes from the  $q_s$  quark component. The gluon-gluon fusion contributes about  $\sim 1\%$  to the total cross section at Tevatron and up to  $\sim 25\%$  at the LHC.

Figs. 2.3 and 2.3 display a comparison between the calculated differential cross sections  $d\sigma/dp_T$  and the experimental data [54–56] at low  $p_T$  ( $p_T < 20$  GeV), and in the full  $p_T$  range. For comparison, we also show the predictions based on the simple  $2 \rightarrow 1$  quark-antiquark annihilation subprocess (dotted histograms), with all quark components summed together. The difference between the results can probably be attributed to the terms not containing large logarithms. The predictions of the “subprocess decomposition” scheme lie by about a factor of 1.25 higher and show better agreement with the data.

Having considered the different partonic subprocesses we see that the dominant contribution comes from the sea quark interactions  $q_s + q_s \rightarrow W/Z$ ,  $q_s + q_v \rightarrow W/Z$  and  $q_s + g \rightarrow W/Z + q'$ . Notably, we find that these subprocesses are mainly due to the quarks emerging from the earlier steps of the parton evolution rather than from the last gluon splitting. Thus, we conclude that the quarks constitute an important component of the parton ladder, not negligible even at the LHC energies and not reducible to the gluon component. Quarks need to be directly included in the evolution equations for consistency and completeness of the latter.

The results of our calculations within the “subprocess decomposition” scheme reasonably agree with the available experimental data and show no need for an extra factor introduced in [57].

## 2.4 Implications for the LHC: Z and W associated with heavy quark pair at Tevatron and the LHC in $k_{\perp}$ -factorization

*Author: Michal Deak, Florian Schwennsen*

To calculate the cross section for  $pp \rightarrow Z/W + Q_i \bar{Q}_j$  with the hard subprocess  $g^* g^* \rightarrow Z/W + Q_i \bar{Q}_j$  at LHC energies we have to convolute the corresponding partonic off-shell cross section with gluon uPDFs. For this purpose we implemented the matrix element squared of the subprocess  $g^* g^* \rightarrow Z/W + Q_i \bar{Q}_j$  into the Monte Carlo generator CASCADE.

Our calculation of the hard matrix elements includes  $W^{\pm}$  and  $Z$  production in association with all possible quark-antiquark channels in gluon gluon fusion. Since the basic structure of all these matrix elements is very similar, we present results only for the typical case of  $Zb\bar{b}$  production at LHC energies of  $\sqrt{s} = 14\text{TeV}$ . We compare our calculation with a prediction using collinear factorization as obtained from the program MCFM [58]. For the collinear factorization calculations we use the parton densities CTEQ6L1 [59]. Since we want compare with NLO collinear calculation, which in MCFM is available only in massless quark approximation, we compare by setting the quark mass to zero in our mass dependent calculation. To emulate the quark mass effect we set a cutoff on the transversal momenta of the quarks with values  $p_{b\perp\min} = m_b = 4.62\text{ GeV}$  in our calculation and in MCFM as well.

The total cross sections are comparable in magnitude, though they differ considerably: 0.406 nb in  $k_{\perp}$ -factorization and 0.748 nb in collinear factorization.

The transverse momentum distribution of the vector boson are shown in Fig. 9. The comparison of the  $k_{\perp}$ -factorization approach to the collinear shows that they agree in transversal momentum distributions of  $Z$  at high values of this quantity. This is no surprise, since at high  $p_{Z\perp}$  the contribution from initial state gluon transverse momenta is expected to become small.

In the distribution of the azimuthal angular distance of  $Z$  and  $\max(p_{b,\perp}, p_{\bar{b},\perp})$  (Fig. 10) we observe that the region from 0 to  $\pi/2$  is forbidden within the collinear calculation due to momentum conservation, which is not the case for  $k_{\perp}$ -factorization. This is caused by the contribution from initial state gluon transversal momentum which allows the transversal momenta of  $Z$ ,  $b$  and  $\bar{b}$  to be unbalanced. A larger spread of possible configurations causes that the distribution in the  $k_{\perp}$ -factorization calculation flattens.

## 3 NLO inclusive jet production in $k_{\perp}$ -factorization

*Author: Jochen Bartels, Agustin Sabio-Vera, Florian Schwennsen*

At different high energy colliders the inclusive jet production is one of the basic measurements. Besides the very successful approach of collinear factorization, also within  $k_{\perp}$ -factorization jet production at HERA has been described [60]. There, the jet vertex has been constructed from the central hard matrix element of quark-antiquark production – connected to the unintegrated gluon and the photon being emitted from the electron. The LO calculation of the corresponding  $\gamma^* g \rightarrow q\bar{q}$  matrix element is straightforward and contains just two diagrams – one sample is shown in Fig. 11a. As it was shown in Ref. [61],  $k_{\perp}$ -factorization in the small  $x$  regime can be formulated at NLO accuracy. However, the jet production at HERA has not been calculated at NLO accuracy so far, but the building blocks are contained in the calculation of the NLO photon impact factor [62–66] including virtual corrections (like in Fig. 11b) and

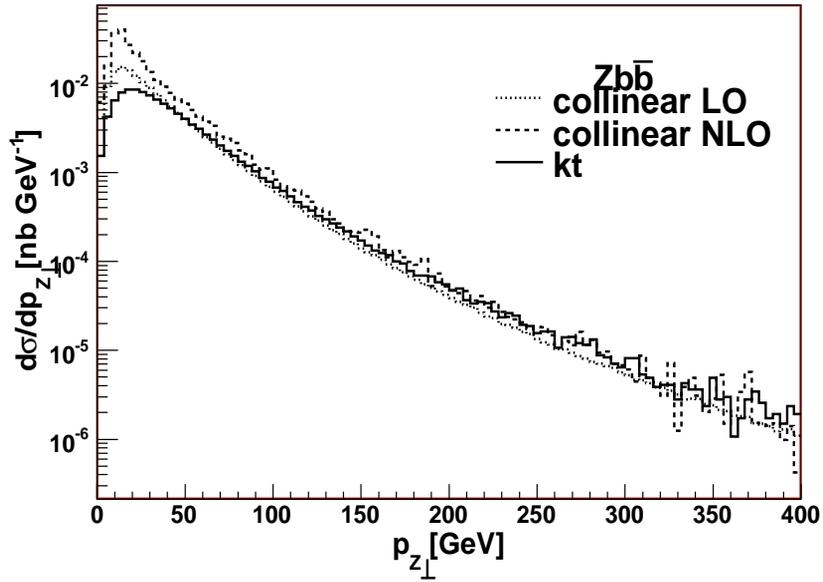


Fig. 9: Comparison of cross sections differential in transverse momentum of the produced  $Z$  gauge boson. Calculation with massless  $b$ -quarks.

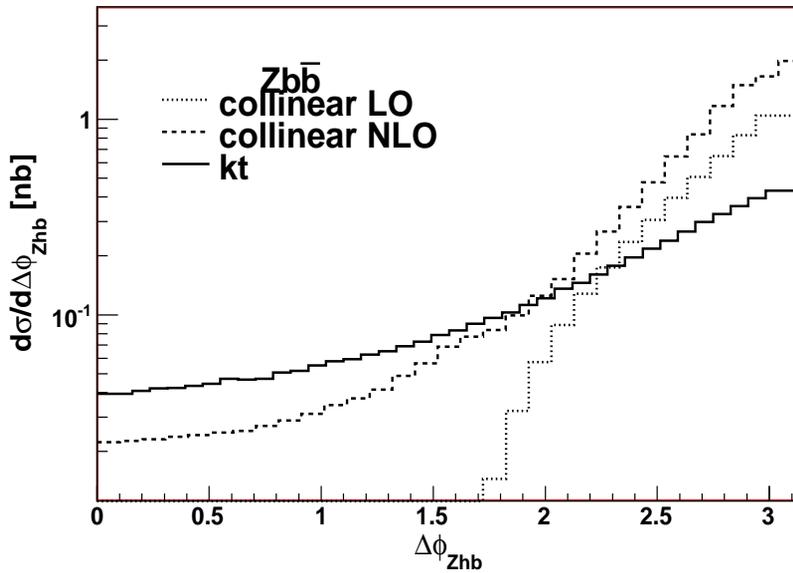


Fig. 10: Comparison of cross sections differential in distance in azimuthal angle of  $Z$  and higher  $p_{\perp} b/\bar{b}$ . Calculation with massless  $b$ -quarks.

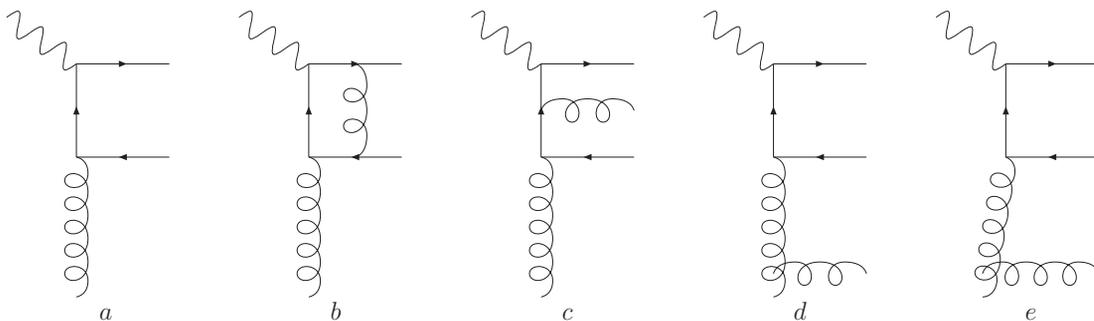


Fig. 11: sample diagrams for the dijet vertex in DIS.

corrections due to the emission of an additional gluon. For such a gluon one has to separate the case where the gluon is ‘close’ to the vertex, giving a standard real correction to the process (like in Fig. 11c), or where the gluon is ‘well separated’ from the vertex (like indicated in Fig. 11d). Another contribution (symbolically indicated in Fig. 11e) would come from the different energy scales at the jet and the proton vertices. The nature of these latter corrections will become more clear when we discuss in the following the jet vertex for hadron-hadron scattering.

In  $k_{\perp}$ -factorization of hadron-hadron collisions, the jet emission vertex can be identified with the Reggeon-Reggeon-gluon vertex (indicated in Fig. 12a). Its square is nothing else but the LO BFKL-kernel. Since the BFKL equation – from which  $k_{\perp}$ -factorization can be derived – has been formulated at NLO [67, 68] as well, it is also possible to calculate the jet emission vertex at this order [61] taking into account that at NLO also the Reggeon-Reggeon-gluon-gluon and Reggeon-Reggeon-quark-antiquark vertices enter the game. It is not sufficient to simply start from the fully integrated emission vertex as used in the NLO BFKL kernel [67, 68]. Rather, one has to carefully separate all the different contributions in their unintegrated form before one can combine them. Moreover, special care has to be taken on the correct treatment of the energy scales involved.

Instead of deriving in detail all the formulas, let us focus on the nature of the different contributions to the NLO jet vertex. In Fig. 12 we represent the different types of contributions by a symbolic diagram. At NLO virtual corrections to the vertex enter the game as shown in Fig. 12b. Since off-shell amplitudes per se are not gauge invariant, the calculation has to be performed as an embedded process. One can *e.g.* consider the process  $q+q \rightarrow q+g+q$  in multi-Regge-kinematics and extract the off-shell Reggeon-Reggeon-gluon vertex. Virtual corrections to this process then also include diagrams (like box diagrams) which do not factorize individually but only in the sum. Therefore, Fig. 12b can only be regarded as one specific example of a virtual correction.

The real corrections to the one jet production consist of two gluon and quark-antiquark production. For the quark-antiquark production one just has to distinguish two cases: either both

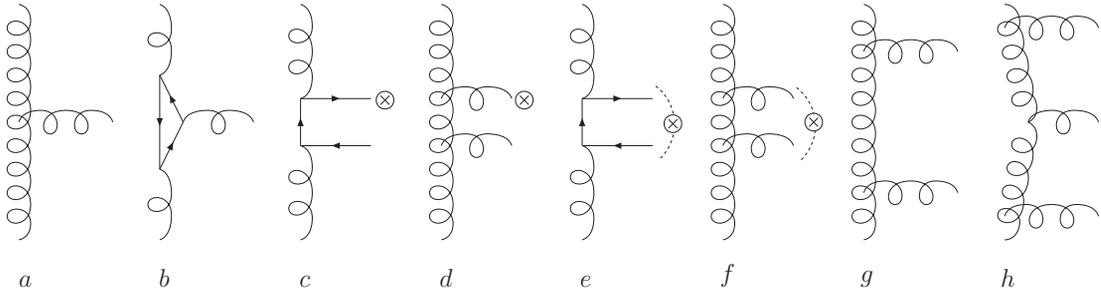


Fig. 12: sample diagrams for the jet vertex in pp.

particles are ‘well separated’ and only one forms the jet while the other contributes to the inclusive part (Fig. 12c), or they are ‘close to each other’ and form one jet (Fig. 12e). The question, whether they are ‘well separated’ or ‘close to each other’, has to be answered by a specific jet definition. The same distinction has to be made for the two gluon production (Fig. 12d and Fig. 12f).

However, the two-gluon production involves some subtleties. Even if the two gluons cannot be combined in one jet, there still remains the question whether the gluon belongs to the same emission vertex or to different ones. A more detailed study of the arrangement of diagrams in the complete framework of NLO BFKL [61, 69] reveals that a contribution from the two neighbored rungs (Fig. 12g) has to be reorganized into the NLO vertex at hand. The contributions in Fig. 12d and Fig. 12g both depend on the scale  $s_\Lambda$  which separates the multi-Regge-kinematics from the quasi-multi-Regge-kinematics. The inclusion of the contribution in Fig. 12g makes the NLO jet vertex (and the NLO BFKL kernel) – to next-to-leading accuracy – independent of  $s_\Lambda$ .

The original formulation of the BFKL approach relies on the scattering of two objects providing an intrinsic and similar hard scale. Instead, in our setting we have to deal with the evolution between the proton – at a soft scale – and the jet – at a hard scale. This imbalance ‘tilts’ the whole evolution such that in fact the BFKL evolution kernel gets a correction introducing some additional collinear evolution. In the language of BFKL this can be understood as the change from a symmetric energy scale  $s_0$  – like  $s_0 = \sqrt{Q_1 Q_2}$  for two colliding objects with scales  $Q_1$  and  $Q_2$  respectively – to an asymmetric one  $s_0 = k_{\perp, \text{Jet}}^2$ . In the end, the result does not depend on this artificial energy scale  $s_0$ , which is ensured by compensating corrections to the impact factors. With respect to the gluon ladder, our jet vertex acts as a kind of impact factor and hence receives from above and below two corrections due to this energy scale change. These complex dependencies involve a large number of concrete diagrams – we only symbolically mark by Fig. 12h that the tilted gluon evolution leads to a correction of the jet emission vertex.

The virtual corrections (Fig. 12b) themselves are infrared divergent. These divergences are canceled by the real corrections (Fig. 12c-f) after they are integrated over. To obtain a jet vertex which is explicitly free of divergences, additional effort has to be made. After identification of those terms in the real corrections which will lead to divergences, one can compensate them by an unintegrated subtraction term, while the integrated subtraction term (which in fact is added such that effectively the result is not changed) cancels the explicit divergences of the virtual

corrections. The exact form of this subtraction term as well as all other formulas which are needed can be found in Ref. [61].

#### 4 Multijet production in the multi-Regge limit: Mueller–Navelet and forward jets

*Authors: Agustín Sabio Vera, Florian Schwennsen*

In this section we briefly describe the calculations performed in [70–73] to obtain the azimuthal angle correlations in Mueller–Navelet jets [74] and forward jets at HERA using the Balitsky–Fadin–Kuraev–Lipatov (BFKL) equation in the next-to-leading (NLO) approximation [67,68] (see also [75–78]). We first comment on the normalized differential cross section for Mueller–Navelet jets. As it is quite insensitive to the parton distribution functions we can simply operate with partonic cross sections, *i.e.*

$$\frac{d\hat{\sigma}}{d^2\vec{q}_1 d^2\vec{q}_2} = \frac{\pi^2 \bar{\alpha}_s^2}{2} \frac{1}{q_1^2 q_2^2} \int \frac{d\omega}{2\pi i} e^{\omega Y} f_{\omega}(\vec{q}_1, \vec{q}_2), \quad (12)$$

where  $\bar{\alpha}_s = \alpha_s N_c / \pi$ ,  $\vec{q}_{1,2}$  are the transverse momenta of the tagged jets, and  $Y$  their relative rapidity. The Green’s function carries the bulk of the  $Y$  dependence and is the solution to the NLO BFKL equation,

$$\left( \omega - \bar{\alpha}_s \hat{K}_0 - \bar{\alpha}_s^2 \hat{K}_1 \right) \hat{f}_{\omega} = \hat{1}, \quad (13)$$

which acts on the basis including the azimuthal angle, *i.e.*,

$$\langle \vec{q} | \nu, n \rangle = \frac{1}{\pi \sqrt{2}} (q^2)^{i\nu - \frac{1}{2}} e^{i n \theta}. \quad (14)$$

As  $Y$  increases the azimuthal angle dependence is controlled by the kernel and it is then reasonable to use LO jet vertices which are much simpler than the NLO ones [79,80]. The differential cross section in the azimuthal angle  $\phi = \theta_1 - \theta_2 - \pi$ , with  $\theta_i$  being the angles of the two tagged jets, reads

$$\frac{d\hat{\sigma}(\alpha_s, Y, p_{1,2}^2)}{d\phi} = \frac{\pi^2 \bar{\alpha}_s^2}{4 \sqrt{p_1^2 p_2^2}} \sum_{n=-\infty}^{\infty} e^{i n \phi} \mathcal{C}_n(Y), \quad (15)$$

where  $p_1$  and  $p_2$  are the cuts on transverse momenta and

$$\mathcal{C}_n(Y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{d\nu}{\left(\frac{1}{4} + \nu^2\right)} \left(\frac{p_1^2}{p_2^2}\right)^{i\nu} e^{\chi(|n|, \frac{1}{2} + i\nu, \bar{\alpha}_s(p_1 p_2)) Y}, \quad (16)$$

and the NLO kernel can be written as

$$\chi(n, \gamma, \bar{\alpha}_s) = \bar{\alpha}_s \chi_0(n, \gamma) + \bar{\alpha}_s^2 \left( \chi_1(n, \gamma) - \frac{\beta_0}{8N_c} \frac{\chi_0(n, \gamma)}{\gamma(1-\gamma)} \right). \quad (17)$$

The eigenvalue of the LO kernel is  $\chi_0(n, \gamma) = 2\psi(1) - \psi(\gamma + \frac{n}{2}) - \psi(1 - \gamma + \frac{n}{2})$ , with  $\psi$  the logarithmic derivative of the Euler function. The action of  $\hat{K}_1$ , in  $\overline{\text{MS}}$  scheme, can be found in [81]. The full cross section only depends on the  $n = 0$  component,

$$\hat{\sigma} = \frac{\pi^3 \bar{\alpha}_s^2}{2 \sqrt{p_1^2 p_2^2}} \mathcal{C}_0(Y). \quad (18)$$

The average of the cosine of the azimuthal angle times an integer projects out the contribution from each of these angular components:

$$\frac{\langle \cos(m\phi) \rangle}{\langle \cos(n\phi) \rangle} = \frac{\mathcal{C}_m(Y)}{\mathcal{C}_n(Y)}. \quad (19)$$

The normalized differential cross section is

$$\frac{1}{\hat{\sigma}} \frac{d\hat{\sigma}}{d\phi} = \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} e^{in\phi} \frac{\mathcal{C}_n(Y)}{\mathcal{C}_0(Y)} = \frac{1}{2\pi} \left\{ 1 + 2 \sum_{n=1}^{\infty} \cos(n\phi) \langle \cos(n\phi) \rangle \right\}. \quad (20)$$

The BFKL resummation is not stable at NLO for zero conformal spin. A manifestation of this lack of convergence is what we found in the gluon–bremsstrahlung scheme where our NLO distributions have an unphysical behavior whenever the  $n = 0$  conformal spin appears in the calculation. To solve this problem we imposed compatibility with renormalization group evolution in the DIS limit following [82–84] for all conformal spins. The new kernel with collinear improvements to all orders in the coupling reads [70–73]

$$\begin{aligned} \omega &= \bar{\alpha}_s (1 + \mathcal{A}_n \bar{\alpha}_s) \left\{ 2\psi(1) - \psi\left(\gamma + \frac{|n|}{2} + \frac{\omega}{2} + \mathcal{B}_n \bar{\alpha}_s\right) \right. \\ &\quad \left. - \psi\left(1 - \gamma + \frac{|n|}{2} + \frac{\omega}{2} + \mathcal{B}_n \bar{\alpha}_s\right) \right\} + \bar{\alpha}_s^2 \left\{ \chi_1(|n|, \gamma) - \frac{\beta_0}{8N_c} \frac{\chi_0(n, \gamma)}{\gamma(1-\gamma)} \right. \\ &\quad \left. - \mathcal{A}_n \chi_0(|n|, \gamma) + \left( \psi'\left(\gamma + \frac{|n|}{2}\right) + \psi'\left(1 - \gamma + \frac{|n|}{2}\right) \right) \left( \frac{\chi_0(|n|, \gamma)}{2} + \mathcal{B}_n \right) \right\}, \quad (21) \end{aligned}$$

where  $\mathcal{A}_n$  and  $\mathcal{B}_n$  are collinear coefficients [70–73]. After this extra resummation our observables have a good physical behavior and are independent of the renormalization scheme. However, it is very important to stress that the asymptotic behavior of the BFKL resummation is convergent for non zero conformal spins. This is why we propose that the ideal distributions to investigate BFKL effects experimentally are those of the form  $\langle \cos(m\phi) \rangle / \langle \cos(n\phi) \rangle$  with  $m, n \neq 0$ , we will see below that in this case the difference between the predictions at LO and at higher orders results is very small.

#### 4.1 Mueller–Navelet jets at the LHC

Long ago, the  $D\emptyset$  [85] collaboration analyzed data for Mueller–Navelet jets at  $\sqrt{s} = 630$  and 1800 GeV. For the angular correlation, LO BFKL predictions were first obtained in [86, 87] and failed to describe the data since the LO results were far too decorrelated. On the other hand, a more conventional fixed order NLO analysis using JETRAD underestimated the decorrelation, while HERWIG was in agreement with the data.

In Fig. 13 we compare the Tevatron data for  $\langle \cos\phi \rangle = \mathcal{C}_1/\mathcal{C}_0$  with our LO, NLO and collinearly resummed predictions. For Tevatron’s cuts, where the lower cut off in transverse momentum for one jet is 20 GeV and for the other 50 GeV, the NLO calculation is unstable under renormalization scheme changes. The convergence of our observables is poor whenever the coefficient associated to zero conformal spin,  $\mathcal{C}_0$ , is used in the calculation. If we eliminate

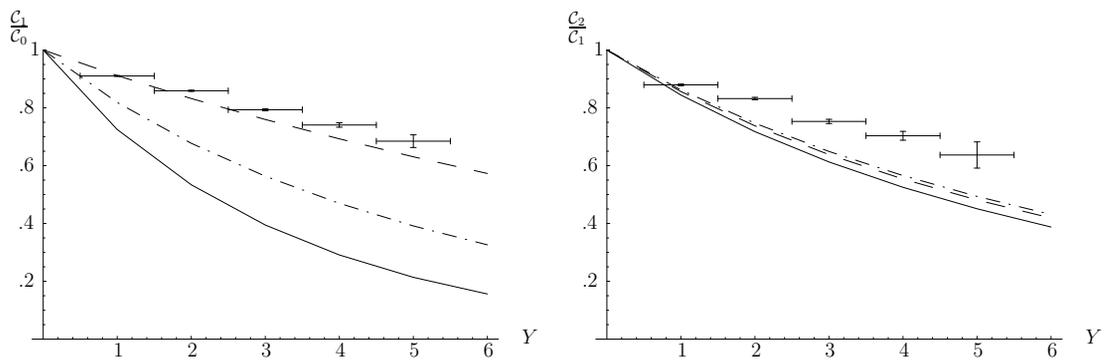


Fig. 13: Left:  $\langle \cos \phi \rangle = C_1/C_0$  and Right:  $\frac{\langle \cos 2\phi \rangle}{\langle \cos \phi \rangle} = \frac{C_2}{C_1}$ , at a  $p\bar{p}$  collider with  $\sqrt{s} = 1.8$  TeV for BFKL at LO (solid) and NLO (dashed). The results from the resummation presented in the text are shown as well (dash-dotted).

this coefficient by calculating the ratios defined in Eq. (19) then the predictions are very stable, see Fig. 13.

The full angular dependence studied at the Tevatron by the  $D0$  collaboration was published in [85]. In Fig. 14 we compare this measurement with the predictions obtained in our approach. For the differential cross section we also make predictions for the LHC at larger  $Y$  in Fig. 15. We estimated several uncertainties in our approach which are represented by gray bands.

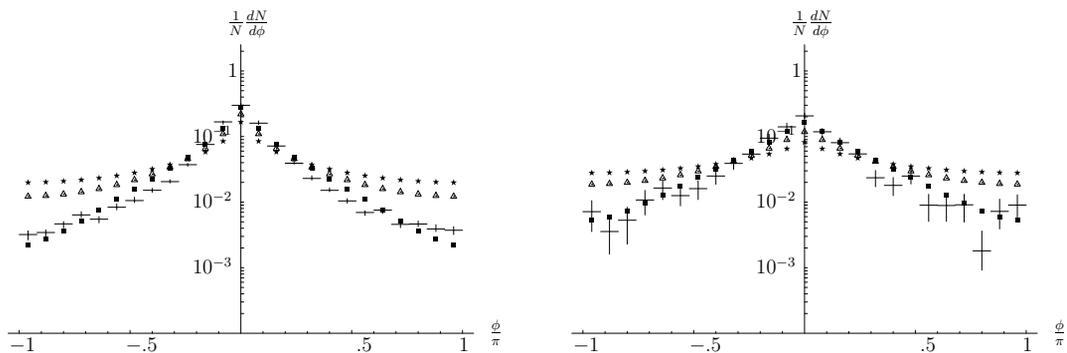


Fig. 14:  $\frac{1}{N} \frac{dN}{d\phi}$  in a  $p\bar{p}$  collider at  $\sqrt{s}=1.8$  TeV using a LO (stars), NLO (squares) and resummed (triangles) BFKL kernel. Plots are shown for  $Y = 3$  (left) and  $Y = 5$  (right).

## 4.2 Forward jets at HERA

In this section we apply the BFKL formalism to predict the decorrelation in azimuthal angle between the electron and a forward jet associated to the proton in Deep Inelastic Scattering (DIS). When the separation in rapidity space between the scattered electron and the forward jet is large and the transverse momentum of the jet is similar to the virtuality of the photon resolving the hadron, then the dominant terms are of BFKL type. This process is similar to that of Mueller–Navelet jets, the only difference being the substitution of one jet vertex by the vertex describing

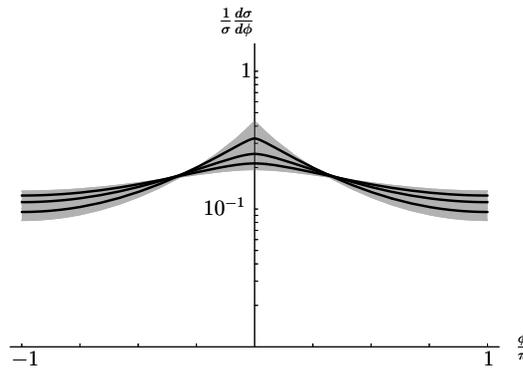


Fig. 15:  $\frac{1}{\sigma} \frac{d\sigma}{d\phi}$  in our resummation scheme for rapidities  $Y = 7, 9, 11$  from top to bottom. The gray band reflects the uncertainty in  $s_0$  and in the renormalization scale  $\mu$ .

the coupling of the electron to the BFKL gluon Green's function via a quark–antiquark pair. Azimuthal angles in forward jets were studied at LO in [88]. We improved their calculation by considering the NLO BFKL kernel and collinear improved versions of it. Fixed order calculations can be found in [89].

In the production of a forward jet in DIS it is necessary to extract a jet with a large longitudinal momentum fraction  $x_{FJ}$  from the proton. When this jet is characterized by a hard scale in the form of a large  $p_t$  it is possible to use conventional collinear factorization to describe the process and the production rate may be written as

$$\sigma(s) = \int dx_{FJ} f_{\text{eff}}(x_{FJ}, \mu_F^2) \hat{\sigma}(\hat{s}), \quad (22)$$

with  $\hat{\sigma}(\hat{s})$  denoting the partonic cross section, and the effective parton density [90] being

$$f_{\text{eff}}(x, \mu_F^2) = G(x, \mu_F^2) + \frac{4}{9} \sum_f [Q_f(x, \mu_F^2) + \bar{Q}_f(x, \mu_F^2)], \quad (23)$$

where the sum runs over all quark flavors, and  $\mu_F$  stands for the factorization scale.

The final expression for the cross section at hadronic level is of the form

$$\frac{d\sigma}{dY d\phi} = C_0(Y) + C_2(Y) \cos 2\phi, \quad (24)$$

with

$$C_n(Y) = \frac{\pi^2 \bar{\alpha}_s^2}{2} \int_{\text{cuts}} dx_{FJ} dQ^2 dy f_{\text{eff}}(x_{FJ}, Q^2) B^{(n)}(y, Q^2, Y) \delta\left(x_{FJ} - \frac{Q^2 e^Y}{ys}\right), \quad (25)$$

where the index in the integral sign refers to the cuts

$$20 \text{ GeV}^2 < Q^2 < 100 \text{ GeV}^2, \quad 0.05 < y < 0.7, \quad 5 \cdot 10^{-3} > x_{Bj} > 4 \cdot 10^{-4}. \quad (26)$$

The integration over the longitudinal momentum fraction  $x_{FJ}$  of the forward jet involves a delta function fixing the rapidity  $Y = \ln x_{FJ}/x_{Bj}$  and  $B^{(n)}$  is a complicated function which can be found in [70–73].

Since the structure of the electron vertex singles out the components with conformal spin 0 and 2, the number of observables related to the azimuthal angle dependence is limited when compared to the Mueller–Navelet case. The most relevant observable is the dependence of the average  $\langle \cos 2\phi \rangle = C_2/C_0$  with the rapidity difference between the forward jet and outgoing lepton. It is natural to expect that the forward jet will be more decorrelated from the leptonic system as the rapidity difference is larger since the phase space for further gluon emission opens up. This is indeed what we observe in our numerical results shown in Fig. 16. We find similar results to the Mueller–Navelet jets case where the most reliable calculation is that with a collinearly-improved kernel. The main effect of the higher order corrections is to increase the azimuthal angle correlation for a given rapidity difference, while keeping the decrease of the correlation as  $Y$  grows.

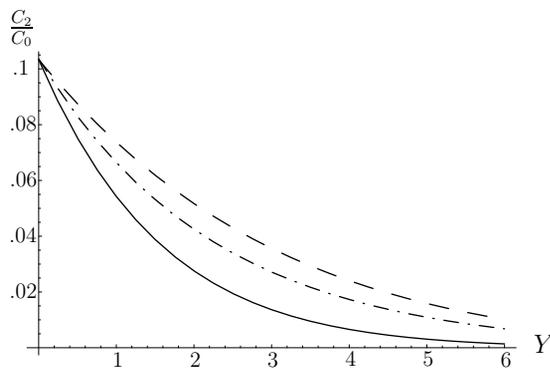


Fig. 16:  $\langle \cos 2\phi \rangle$  at the  $ep$  collider HERA at leading (solid), next to leading order (dashed), and for resummed kernel (dash-dotted).

## 5 NLL BFKL effects: Mueller-Navelet and forward jets

*Author: Christophe Royon*

### 5.1 Forward jets at HERA

Following the successful BFKL [91–93] parametrisation of the forward-jet cross-section  $d\sigma/dx$  at Leading Order (LO) at HERA [75, 94, 95], it is possible to perform a similar study using Next-to-leading (NLL) resummed BFKL kernels. This method can be used for forward jet production at HERA in particular, provided one takes into account the right two scales of the forward-jet problem, namely  $Q^2$  for the lepton and  $k_T^2$  for the jet vertex respectively. In this short report, we will only discuss the phenomenological aspects and all detailed calculations can be found in Ref. [71, 73, 76, 77] for forward jets at HERA and in Ref. [78] for Mueller-Navelet jets at the Tevatron and the LHC.

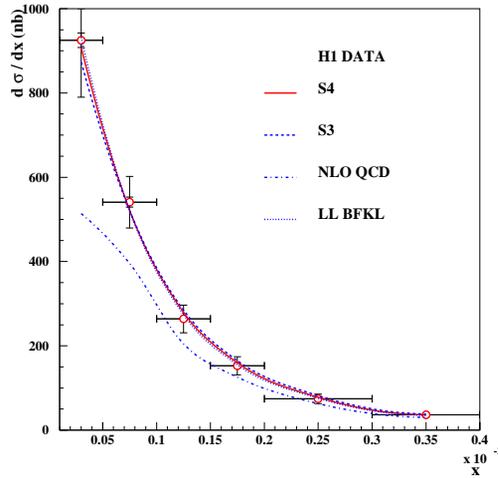


Fig. 17: Comparison between the H1  $d\sigma/dx$  measurement with predictions for BFKL-LL, BFKL-NLL (S3 and S4 schemes) and DGLAP NLO calculations (see text). S4, S3 and LL BFKL cannot be distinguished on that figure.

The BFKL NLL [67, 68, 96] longitudinal transverse cross section reads:

$$\frac{d\sigma_{T,L}^{\gamma^*p \rightarrow JX}}{dx_J dk_T^2} = \frac{\alpha_s(k_T^2)\alpha_s(Q^2)}{k_T^2 Q^2} f_{eff}(x_J, k_T^2) \int d\gamma \left(\frac{Q^2}{k_T^2}\right)^\gamma \phi_{T,L}^\gamma(\gamma) e^{\bar{\alpha}(k_T Q)\chi_{eff}[\gamma, \bar{\alpha}(k_T Q)]Y} \quad (27)$$

where the  $\chi_{eff}$  is the effective BFKL NLL kernel and the  $\phi$  are the transverse and longitudinal impact factors taken at LL. The effective kernel  $\chi_{eff}(\gamma, \bar{\alpha})$  is defined from the NLL kernel  $\chi_{NLL}(\gamma, \omega)$  by solving the implicit equation numerically

$$\chi_{eff}(\gamma, \bar{\alpha}) = \chi_{NLL}[\gamma, \bar{\alpha} \chi_{eff}(\gamma, \bar{\alpha})] . \quad (28)$$

The integration over  $\gamma$  in Eq. 27 is performed numerically. It is possible to fit directly  $d\sigma/dx$  measured [97] by the H1 collaboration using this formalism with one single parameter, the normalisation. The values of  $\chi_{NLL}$  are taken at NLL [67, 68, 96] using different resummation schemes to remove spurious singularities defined as S3 and S4 [82]. Contrary to LL BFKL, it is worth noticing that the coupling constant  $\alpha_s$  is taken using the renormalisation group equations, the only free parameter in the fit being the normalisation.

To compute  $d\sigma/dx$  in the experimental bins, we need to integrate the differential cross section on the bin size in  $Q^2$ ,  $x_J$  (the momentum fraction of the proton carried by the forward jet),  $k_T$  (the jet transverse momentum), while taking into account the experimental cuts. To simplify the numerical calculation, we perform the integration on the bin using the variables where the cross section does not change rapidly, namely  $k_T^2/Q^2$ ,  $\log 1/x_J$ , and  $1/Q^2$ . Experimental cuts are treated directly at the integral level (the cut on  $0.5 < k_T^2/Q^2 < 5$  for instance) or using a toy Monte Carlo. More detail can be found about the fitting procedure in Appendix A of Ref. [75].

The NLL fits [71, 73, 76, 77] can nicely describe the H1 data [97] for the S4 and S3 schemes [71, 73, 75–77, 94, 95] ( $\chi^2 = 0.48/5$  and  $\chi^2 = 1.15/5$  respectively per degree of

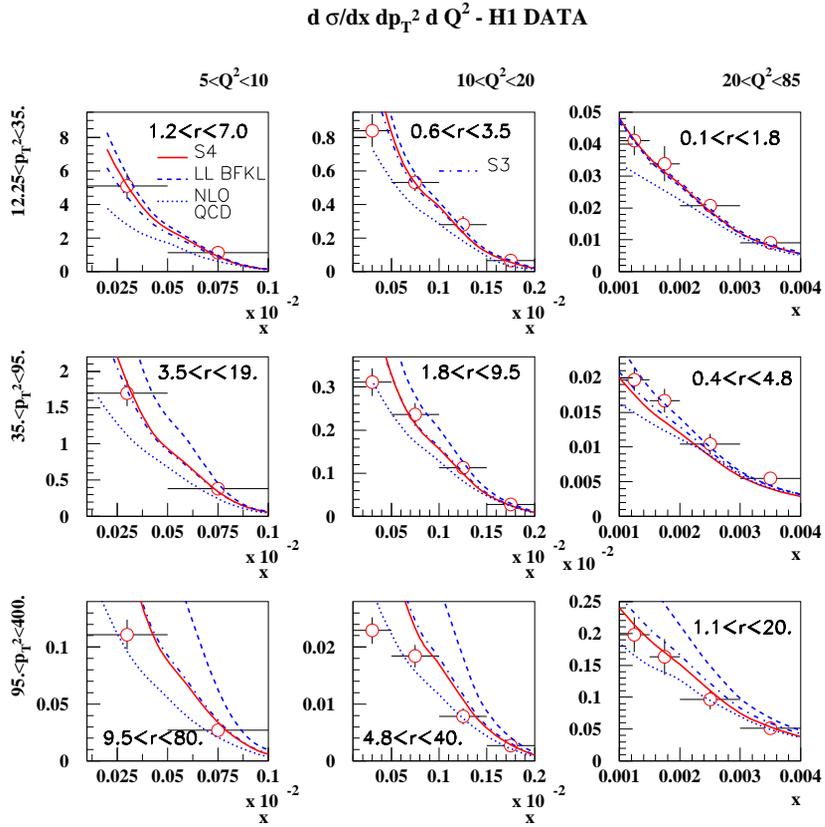


Fig. 18: Comparison between the H1 measurement of the triple differential cross section with predictions for BFKL-LL, BFKL-NLL and DGLAP NLO calculations (see text).

freedom with statistical and systematic errors added in quadrature). The curve using a LL fit is indistinguishable in Fig. 17 from the result of the BFKL-NLL fit. The DGLAP NLO calculation fails to describe the H1 data at lowest  $x$  (see Fig. 17). We also checked the effect of changing the scale in the exponential of Eq. 27 from  $k_T Q$  to  $2k_T Q$  or  $k_T Q/2$  which leads to a difference of 20% on the cross section while changing the scale to  $k_T^2$  or  $Q^2$  modifies the result by less than 5% which is due to the cut on  $0.5 < k_T^2/Q^2 < 5$ . Implementing the higher-order corrections in the impact factor due to exact gluon dynamics in the  $\gamma^* \rightarrow q\bar{q}$  transition [98] changes the result by less than 3%.

The H1 collaboration also measured the forward jet triple differential cross section [97] and the results are given in Fig. 18. We keep the same normalisation coming from the fit to  $d\sigma/dx$  to predict the triple differential cross section. The BFKL LL formalism leads to a good description of the data when  $r = k_T^2/Q^2$  is close to 1 and deviates from the data when  $r$  is further away from 1. This effect is expected since DGLAP radiation effects are supposed to occur when the ratio between the jet  $k_T$  and the virtual photon  $Q^2$  are further away from 1. The BFKL NLL calculation including the  $Q^2$  evolution via the renormalisation group equation leads to a good description of the H1 data on the full range. We note that the higher order corrections are small when  $r \sim 1$ , when the BFKL effects are supposed to dominate. By contrast, they are significant as expected when  $r$  is different from one, i.e. when DGLAP evolution becomes relevant. We notice that the DGLAP NLO calculation fails to describe the data when  $r \sim 1$ , or in the region where BFKL resummation effects are expected to appear.

In addition, we checked the dependence of our results on the scale taken in the exponential of Eq. 27. The effect is a change of the cross section of about 20% at low  $p_T$  increasing to 70% at highest  $p_T$ . Taking the correct gluon kinematics in the impact factor lead as expected to a better description of the data at high  $p_T$  [71, 73, 76, 77].

## 5.2 Mueller-Navelet jets at the Tevatron and the LHC

Mueller-Navelet jets are ideal processes to study BFKL resummation effects [74]. Two jets with a large interval in rapidity and with similar transverse momenta are considered. A typical observable to look for BFKL effects is the measurement of the azimuthal correlations between both jets. The DGLAP prediction is that this distribution should peak towards  $\pi$  - i.e. jets are back-to-back- whereas multi-gluon emission via the BFKL mechanism leads to a smoother distribution. The relevant variables to look for azimuthal correlations are the following:

$$\begin{aligned}\Delta\eta &= y_1 - y_2 \\ y &= (y_1 + y_2)/2 \\ Q &= \sqrt{k_1 k_2} \\ R &= k_2/k_1\end{aligned}$$

The azimuthal correlation for BFKL reads:

$$2\pi \frac{d\sigma}{d\Delta\eta dR d\Delta\Phi} \bigg/ \frac{d\sigma}{d\Delta\eta dR} = 1 + \frac{2}{\sigma_0(\Delta\eta, R)} \sum_{p=1}^{\infty} \sigma_p(\Delta\eta, R) \cos(p\Delta\Phi)$$

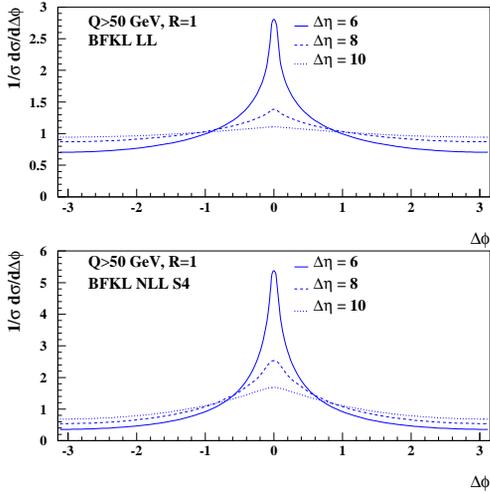


Fig. 19: The Mueller-Navelet jet  $\Delta\Phi$  distribution for LHC kinematics in the BFKL framework at LL (upper plots) and NLL-S4 (lower plots) accuracy for  $\Delta\eta = 6, 8, 10$ .

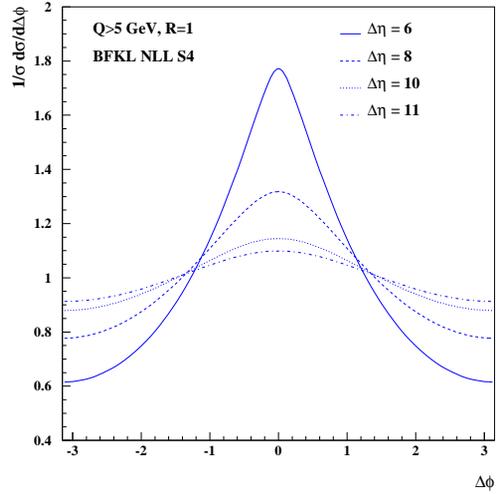


Fig. 20: Azimuthal correlations between jets with  $\Delta\eta = 6, 8, 10$  and  $11$  and  $p_T > 5$  GeV in the CDF acceptance. This measurement will represent a clear test of the BFKL regime.

where in the NLL BFKL framework,

$$\sigma_p = \int_{E_T}^{\infty} \frac{dQ}{Q^3} \alpha_s(Q^2/R) \alpha_s(Q^2 R) \left( \int_{y_<}^{y_>} dy x_1 f_{eff}(x_1, Q^2/R) x_2 f_{eff}(x_2, Q^2 R) \right) \int_{1/2-\infty}^{1/2+\infty} \frac{d\gamma}{2i\pi} R^{-2\gamma} e^{\bar{\alpha}(Q^2) \chi_{eff}(p, \gamma, \bar{\alpha}) \Delta\eta}$$

and  $\chi_{eff}$  is the effective resummed kernel. Computing the different  $\sigma_p$  at NLL for the resummation schemes S3 and S4 allowed us to compute the azimuthal correlations at NLL. As expected, the  $\Delta\Phi$  dependence is less flat than for BFKL LL and is closer to the DGLAP behaviour [78]. In Fig. 19, we display the observable  $1/\sigma d\sigma/d\Delta\Phi$  as a function of  $\Delta\Phi$ , for LHC kinematics. The results are displayed for different values of  $\Delta\eta$  and at both LL and NLL accuracy using the S4 resummation scheme. In general, the  $\Delta\Phi$  spectra are peaked around  $\Delta\Phi = 0$ , which is indicative of jet emissions occurring back-to-back. In addition the  $\Delta\Phi$  distribution flattens with increasing  $\Delta\eta = y_1 - y_2$ . Note the change of scale on the vertical axis which indicates the magnitude of the NLL corrections with respect to the LL-BFKL results. The NLL corrections slow down the azimuthal angle decorrelations for both increasing  $\Delta\eta$  and  $R$  deviating from 1. We also studied the  $R$  dependence of our prediction which is quite weak [78]. We also studied the scale dependence of our results by modifying the scale  $Q^2$  to either  $Q^2/2$  or  $2Q^2$  and the effect on the azimuthal distribution is of the order of 20%. The effect of the energy conservation in the BFKL equation [78] is large when  $R$  goes away from 1. The effect is to reduce the effective value of  $\Delta\eta$  between the jets and thus the decorrelation effect. However, it is worth noticing that this effect is negligible when  $R$  is close to 1 where this measurement will be performed.

A measurement of the cross-section  $d\sigma^{hh \rightarrow JXJ}/d\Delta\eta dR d\Delta\Phi$  at the Tevatron (Run 2) or

the LHC will allow for a detailed study of the BFKL QCD dynamics since the DGLAP evolution leads to much less jet angular decorrelation (jets are back-to-back when  $R$  is close to 1). In particular, measurements with values of  $\Delta\eta$  reaching 8 or 10 will be of great interest, as these could allow to distinguish between BFKL and DGLAP resummation effects and would provide important tests for the relevance of the BFKL formalism.

To illustrate this result, we give in Fig. 20 the azimuthal correlation in the CDF acceptance. The CDF collaboration installed the mini-Plugs calorimeters aiming for rapidity gap selections in the very forward regions and these detectors can be used to tag very forward jets. A measurement of jet  $p_T$  with these detectors would not be possible but their azimuthal segmentation allows a  $\phi$  measurement. In Fig. 20, we display the jet azimuthal correlations for jets with a  $p_T > 5$  GeV and  $\Delta\eta = 6, 8, 10$  and 11. For  $\Delta\eta = 11$ , we notice that the distribution is quite flat, which would be a clear test of the BFKL prediction.

## 6 Forward Jets in the CASTOR calorimeter in the CMS experiment

*Author: Albert Knutsson*

The CASTOR (Centauro and STRange Object Research) detector [99] is a Cherenkov radiation calorimeter consisting of tungsten absorber plates sandwiched with plates of quartz, used as the active material in the detector. The construction is repeated in octants in azimuthal angle, giving a full  $360^\circ$  coverage. For each octant lightguides and photomultipliers are situated on top of the plates in two columns, 14 channels deep along the beam direction. 2 of the 14 channels are designed for detection of electromagnetic particles and the rest are hadronic channels. Thus the detector consists of a total number of  $16 \times 14$  channels. CASTOR is situated 14.4 m from the interaction vertex in the CMS detector at LHC and covers the pseudorapidity range  $5.2 < \eta < 6.6$ .

Since CASTOR has no segmentation in polar angle it will not be possible to define jets according to conventional jet algorithms which use the energy, polar and azimuthal angle of particles. Here we investigate the possibility to measure jet events with CASTOR, by using only the azimuthal segmentation and energy deposition. The studies are carried out on Monte Carlo generator level. Events are generated with the ARIADNE event generator [100], with the hadron level jets defined according to the inclusive  $k_T$  algorithm. The kinematic region has been divided into 16 slices in phi, for which the energies of all particles are summed. In Fig. 21a-b the correlations between the hadron level jet energy and different energy depositions in the CASTOR region are shown. Clearly the total energy contained in the CASTOR region is too large compared to the energy of the hadron level jet (Fig. 21a), while the energy in the phi segment with highest energy gives a better correlation with the true jet energy (Fig. 21b). The best reconstruction of the jet energy is achieved if the energy in the most active phi segment is summed with the two neighbouring cells (Fig. 21c). This is the method we use in the physics studies presented in the next section. In future, one can improve the jet reconstruction with more complex algorithms in order to obtain an even better correlation in jet energy. Finally, in Fig. 22 the azimuthal angle of the jet axis is plotted versus the azimuthal slice with highest energy. A good correlation is seen.

Events in which an energetic jet is produced close to the proton remnant (the forward direction) are sensitive to the higher order reactions due to the long rapidity range available for

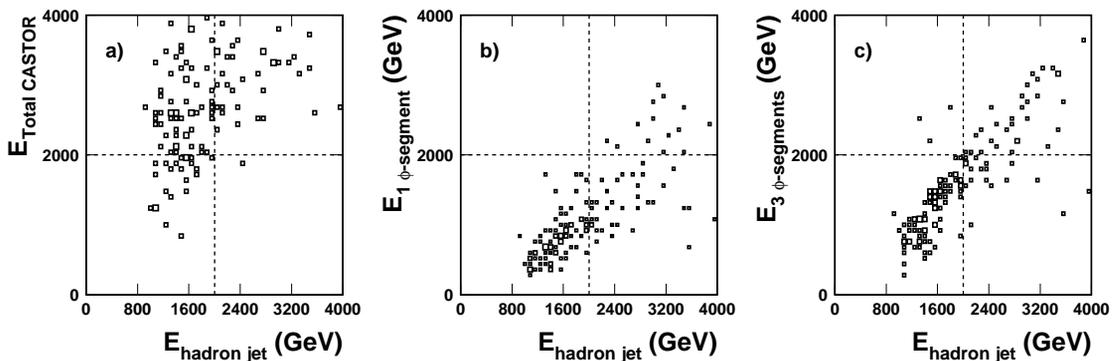


Fig. 21: Monte Carlo predictions on generator level for the hadron level jet energy vs. **a** the energy contained in the full CASTOR region, **b** the energy in the  $\phi$ -segment with the highest energy deposit and **c** the energy in the  $\phi$ -segment with the highest energy deposit summed with the energy in two neighbouring cells.

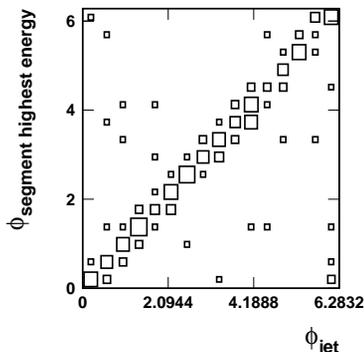


Fig. 22: Azimuthal angle of the highest energy  $\phi$ -segment vs. the azimuthal angle of the jet.

radiation between the jet and the hard scattering vertex. The longitudinal momentum fraction of the proton,  $x$ , can be related to the rapidity,  $y$ , by approximately  $x \sim e^{-y}$ , which further suggests that forward physics gives us valuable information about low  $x$  parton dynamics.

At HERA, forward jet events have been analysed [97, 101] and improved our understanding of QCD. Available fixed order calculations (next-to-leading order  $O(\alpha_s^2)$ ) as well as the higher order reactions approximated by DGLAP parton showers underestimate the HERA data by up to a factor of 2. The data can be described only if the ordering of the transverse momenta of the radiated gluons is broken in the theoretical predictions.

In events where the transverse momentum of the forward jet is close to the scale of the hard reaction the DGLAP like scenario, i.e. events with QCD radiation ordered in transverse momenta, is further suppressed. In the HERA analysis this is achieved by requiring that the square of the transverse momentum of the forward jet was in the same order as the virtuality of the exchanged photon. In the analysis presented here we instead require that two additional hard jets are produced in the central region of the detector. For the forward jets in the CASTOR region this gives up to 5 units of pseudorapidity range available for more gluon radiation.

For this study the Monte Carlo events are generated by using the full event generators PYTHIA 6.4.14 [9] and ARIADNE 1.4 [100]. PYTHIA 6.4.14 is based on LO DGLAP parton showers, which gives gluon radiation ordered in transverse momentum with respect to rapidity. In ARIADNE, parton showers are generated by the Color Dipole Model (CDM), resulting in gluon radiation without any ordering in transverse momentum with respect to rapidity. This corresponds to a BFKL like final state. PYTHIA is run with the so called tune A multiple interaction model.

The hadron level jets are defined with the inclusive  $k_t$  algorithm. Events are selected which contain a jet with a transverse momentum  $E_T > 10$  GeV and a pseudorapidity  $5.2 < \eta < 6.6$ . To further suppress events with DGLAP like dynamics, two jets with  $E_T > 10$  GeV are required in the central region,  $|\eta| < 1.5$ . The resulting cross-section is shown in Fig. 23 as a function of the forward jet energy. As can be seen, CDM is producing more jets at higher energies, while the events with gluon emissions generated according to DGLAP dynamics have a suppressed jet production. At the highest forward jet energies the difference between the models is up to two orders of magnitude.

In the following we replace the hadronic forward jet with the jet reconstruction described in the previous section. In addition we have also smeared the particle energies according to resolutions measured in the CASTOR beam test [102] and applied a noise cut at 1 GeV. Since we can no longer determine the  $E_T$  of the reconstructed forward jet, this cut is removed from the forward jet selection, but the measured range in energy is kept. The major consequence is, as expected, an increased number of jets at low energies. The predictions from PYTHIA and CDM shows that the very high sensitivity to the scheme used for the QCD radiation is still preserved (see Fig. 24).

In Fig. 25a and b we investigate the PDF uncertainties for the suggested measurement. As can be seen in Fig. 25a the predicted forward jet cross section does not distinguish between PDFs which has been fitted at leading order with LO  $\alpha_s$ , CTEQ6L, or NLO  $\alpha_s$ , CTEQ6LL. Using the CTEQ6.5 PDF however gives a lower forward jet cross section, as illustrated in Fig. 25b. Here the PDF uncertainty, based on the 40 error eigensets for CTEQ6.5, are shown for the PYTHIA prediction.

Finally, the response to multiple interactions (MI) is studied in Fig. 26. We see that the impact of MI is expected to be large in general; excluding MI lowers the cross section by roughly an order of magnitude. Except of that, the sensitivity to the different MI tunes and models are fairly small in comparison to the impact of using a CDM.

In summary a method for jet reconstruction in the CASTOR calorimeter has been suggested. We have shown that using the method to measure forward jets in CASTOR in addition to two jets in the central region may give a very large sensitivity to the dynamics of the parton shower. This is also true if PDF uncertainties and different MI models are taken into account.

## References

- [1] NLO Multileg Working Group Collaboration, Z. Bern *et al.* (2008), arXiv:0803.0494 [hep-ph].

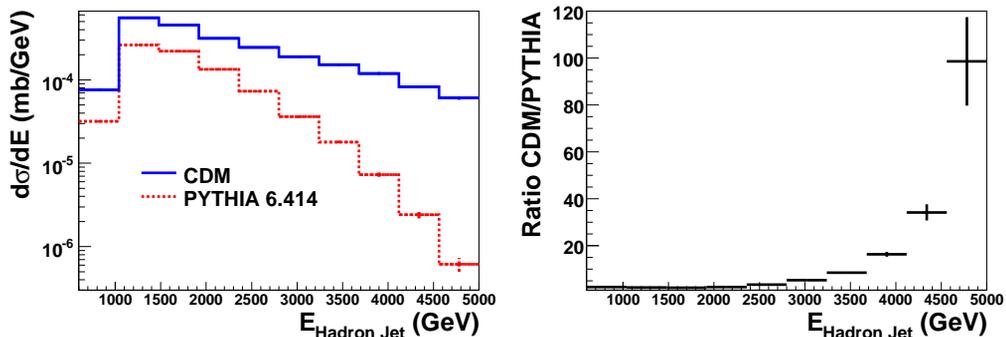


Fig. 23: Hadron level cross sections for events with two central jets and a forward jet in the pseudorapidity region of the CASTOR calorimeter. The predictions are on generator level without any assumptions about the detector.

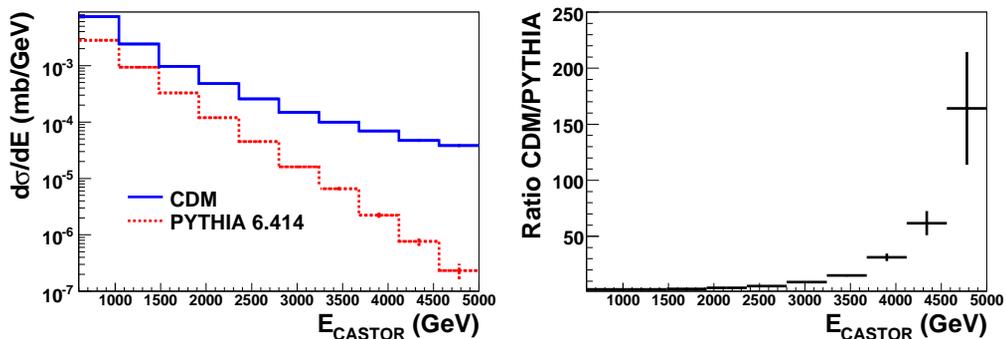


Fig. 24: Cross sections for events with two central jets and a forward jet in the pseudorapidity region of the CASTOR calorimeter. The predictions are on generator level, but here with the forward jet reconstructed as described in the text and forward particle momenta smeared according to CASTOR beam test data.

- [2] S. Hoche *et al.* (2006). hep-ph/0602031.
- [3] N. Lavesson and L. Lonnblad, JHEP **04**, 085 (2008), arXiv:0712.2966 [hep-ph].
- [4] J. Alwall *et al.*, Eur. Phys. J. **C53**, 473 (2008).
- [5] S. Alekhin *et al.* (2005), arXiv:hep-ph/0601012.
- [6] S. Alekhin *et al.* (2005), arXiv:hep-ph/0601013.
- [7] G. Corcella *et al.*, JHEP **01**, 010 (2001), arXiv:hep-ph/0011363.
- [8] G. Corcella *et al.* (2002), arXiv:hep-ph/0210213.
- [9] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006), arXiv:hep-ph/0603175.
- [10] G. Marchesini and B. R. Webber, Nucl. Phys. **B386**, 215 (1992).

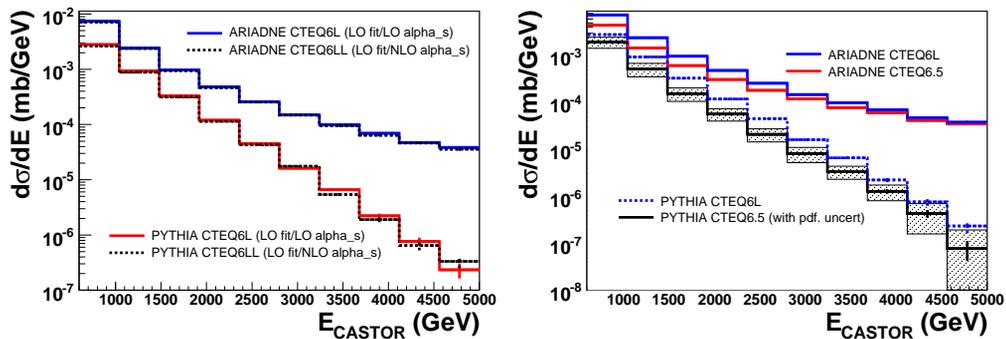


Fig. 25: The PDF uncertainties for the 2+forward jet cross section. The predictions are on generator level, but with the forward jet reconstructed as described in the text and forward particle momenta smeared according to CASTOR beam test data.

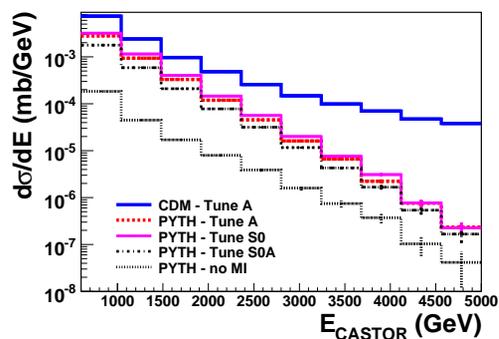


Fig. 26: Monte Carlo prediction for the 2+forward jet cross section using different MI models and tunes. The predictions are on generator level, but with the forward jet reconstructed as described in the text and forward particle momenta smeared according to CASTOR beam test data.

- [11] M. Ciafaloni, Nucl. Phys. **B296**, 49 (1988).
- [12] S. Catani, M. Ciafaloni, and F. Hautmann, Phys. Lett. **B242**, 97 (1990).
- [13] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).
- [14] S. Catani, M. Ciafaloni, and F. Hautmann, Phys. Lett. **B307**, 147 (1993).
- [15] Small x Collaboration, B. Andersson *et al.*, Eur. Phys. J. **C25**, 77 (2002), arXiv:hep-ph/0204115.
- [16] H. Jung, Comput. Phys. Commun. **143**, 100 (2002), arXiv:hep-ph/0109102.
- [17] H. Jung and G. P. Salam, Eur. Phys. J. **C19**, 351 (2001), arXiv:hep-ph/0012143.
- [18] L. Lonnblad and M. Sjoedahl, JHEP **05**, 038 (2005), arXiv:hep-ph/0412111.
- [19] L. Lonnblad and M. Sjoedahl, JHEP **02**, 042 (2004), arXiv:hep-ph/0311252.

- [20] G. Gustafson, L. Lonnblad, and G. Miu, *JHEP* **09**, 005 (2002), arXiv:hep-ph/0206195.
- [21] K. J. Golec-Biernat, S. Jadach, W. Placzek, P. Stephens, and M. Skrzypek, *Acta Phys. Polon.* **B38**, 3149 (2007), arXiv:hep-ph/0703317.
- [22] S. Hoche, F. Krauss, and T. Teubner, *Multijet events in the  $k_T$ -factorisation scheme*, 2007. arXiv:0705.4577 [hep-ph].
- [23] Small x Collaboration, J. R. Andersen *et al.*, *Eur. Phys. J.* **C35**, 67 (2004), arXiv:hep-ph/0312333.
- [24] F. Hautmann and H. Jung (2007), arXiv:0712.0568 [hep-ph].
- [25] F. Hautmann and H. Jung, *JHEP* **10**, 113 (2008), arXiv:0805.1049 [hep-ph].
- [26] M. Ciafaloni, *PoS RADCOR2007*, 029 (2007).
- [27] G. Altarelli, R. D. Ball, and S. Forte, *PoS RADCOR2007*, 028 (2007), arXiv:0802.0968 [hep-ph].
- [28] J. Collins, *Rapidity divergences and valid definitions of parton densities*, 2008. arXiv:0808.2665 [hep-ph].
- [29] J. C. Collins (2001), arXiv:hep-ph/0106126.
- [30] S. Catani and F. Hautmann, *Nucl. Phys.* **B427**, 475 (1994), arXiv:hep-ph/9405388.
- [31] S. Catani and F. Hautmann, *Phys. Lett.* **B315**, 157 (1993).
- [32] L. N. Lipatov, *Phys. Rept.* **286**, 131 (1997), arXiv:hep-ph/9610276.
- [33] M. Buza, Y. Matiounine, J. Smith, R. Migneron, and W. L. van Neerven, *Nucl. Phys.* **B472**, 611 (1996), arXiv:hep-ph/9601302.
- [34] S. Riemersma, J. Smith, and W. L. van Neerven, *Phys. Lett.* **B347**, 143 (1995), arXiv:hep-ph/9411431.
- [35] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, *Nucl. Phys.* **B392**, 162 (1993).
- [36] S. Moch, J. A. M. Vermaseren, and A. Vogt, *Phys. Lett.* **B606**, 123 (2005), arXiv:hep-ph/0411112.
- [37] A. Vogt, S. Moch, and J. A. M. Vermaseren, *Nucl. Phys.* **B691**, 129 (2004), arXiv:hep-ph/0404111.
- [38] DELPHI Collaboration, P. Abreu *et al.*, *Z. Phys.* **C73**, 11 (1996).
- [39] H1 Collaboration, C. Adloff *et al.*, *Eur. Phys. J.* **C21**, 33 (2001), hep-ex/0012053.

- [40] W. Press *et al.*, *Numerical Recipes*. Cambridge University Press, 1992.
- [41] F. James and M. Roos, *Comput. Phys. Commun.* **10**, 343 (1975).
- [42] J. C. Collins and R. K. Ellis, *Nucl. Phys.* **B360**, 3 (1991).
- [43] S. Berge, P. M. Nadolsky, F. Olness, and C. P. Yuan, *Phys. Rev.* **D72**, 033015 (2005), [arXiv:hep-ph/0410375](#).
- [44] M. Deák and F. Schwennsen, *JHEP* **09**, 035 (2008), [arXiv:0805.3763 \[hep-ph\]](#).
- [45] S. P. Baranov, A. V. Lipatov, and N. P. Zotov, *Phys. Rev.* **D78**, 014025 (2008), [arXiv:0805.4821 \[hep-ph\]](#).
- [46] S. P. Baranov, A. V. Lipatov, and N. P. Zotov, *Yad. Fiz.* **67**, 856 (2004), [arXiv:hep-ph/0302171](#).
- [47] A. V. Lipatov and N. P. Zotov, *Phys. Rev.* **D75**, 014028 (2007), [arXiv:hep-ph/0611302](#).
- [48] F. Hautmann, *Phys. Lett.* **B535**, 159 (2002), [arXiv:hep-ph/0203140](#).
- [49] A. V. Lipatov and N. P. Zotov, *Eur. Phys. J.* **C44**, 559 (2005), [arXiv:hep-ph/0501172](#).
- [50] M. A. Kimber, A. D. Martin, and M. G. Ryskin, *Phys. Rev.* **D63**, 114027 (2001), [hep-ph/0101348](#).
- [51] G. Watt, A. D. Martin, and M. G. Ryskin, *Eur. Phys. J.* **C31**, 73 (2003), [arXiv:hep-ph/0306169](#).
- [52] S. P. Baranov, A. V. Lipatov, and N. P. Zotov, *Phys. Rev.* **D77**, 074024 (2008), [arXiv:0708.3560 \[hep-ph\]](#).
- [53] A. Kulesza and W. J. Stirling, *Nucl. Phys.* **B555**, 279 (1999), [arXiv:hep-ph/9902234](#).
- [54] D0 Collaboration, B. Abbott *et al.*, *Phys. Rev.* **D61**, 032004 (2000), [arXiv:hep-ex/9907009](#).
- [55] D0 Collaboration, B. Abbott *et al.*, *Phys. Lett.* **B513**, 292 (2001), [arXiv:hep-ex/0010026](#).
- [56] CDF Collaboration, A. A. Affolder *et al.*, *Phys. Rev. Lett.* **84**, 845 (2000), [arXiv:hep-ex/0001021](#).
- [57] G. Watt, A. D. Martin, and M. G. Ryskin, *Phys. Rev.* **D70**, 014012 (2004), [arXiv:hep-ph/0309096](#).

- [58] J. M. Campbell, R. K. Ellis, and D. L. Rainwater, Phys. Rev. **D68**, 094021 (2003),  
arXiv:hep-ph/0308195;  
J. M. Campbell and R. K. Ellis, Phys. Rev. **D65**, 113007 (2002),  
arXiv:hep-ph/0202176;  
J. Campbell and K. Ellis <http://mcfm.fnal.gov/>.
- [59] J. Pumplin, D. Stump, J. Huston, H. Lai, P. Nadolsky, and W. Tung, JHEP **0207**, 012 (2002). <http://www.citebase.org/abstract?id=oai:arXiv.org:hep-ph/0201195>.
- [60] S. P. Baranov and N. P. Zotov, Phys. Lett. **B491**, 111 (2000).
- [61] J. Bartels, A. Sabio Vera, and F. Schwennsen, JHEP **0611**, 051 (2006),  
hep-ph/0608154.
- [62] J. Bartels, S. Gieseke, and C. F. Qiao, Phys. Rev. **D63**, 056014 (2001),  
arXiv:hep-ph/0009102.
- [63] J. Bartels, S. Gieseke, and A. Kyrieleis, Phys. Rev. **D65**, 014006 (2002),  
arXiv:hep-ph/0107152.
- [64] J. Bartels, D. Colferai, S. Gieseke, and A. Kyrieleis, Phys. Rev. **D66**, 094017 (2002),  
arXiv:hep-ph/0208130.
- [65] J. Bartels and A. Kyrieleis, Phys. Rev. **D70**, 114003 (2004),  
arXiv:hep-ph/0407051.
- [66] G. Chachamis and J. Bartels, PoS **DIFF2006**, 026 (2006).
- [67] V. S. Fadin and L. N. Lipatov, Phys. Lett. **B429**, 127 (1998),  
arXiv:hep-ph/9802290.
- [68] M. Ciafaloni and G. Camici, Phys. Lett. **B430**, 349 (1998),  
arXiv:hep-ph/9803389.
- [69] V. S. Fadin (1998), arXiv:hep-ph/9807528.
- [70] A. Sabio Vera, Nucl. Phys. **B746**, 1 (2006), arXiv:hep-ph/0602250.
- [71] A. Sabio Vera and F. Schwennsen, Nucl. Phys. **B776**, 170 (2007),  
arXiv:hep-ph/0702158.
- [72] F. Schwennsen (2007), arXiv:hep-ph/0703198.
- [73] A. Sabio Vera and F. Schwennsen, Phys. Rev. **D77**, 014001 (2008),  
arXiv:0708.0549 [hep-ph].
- [74] A. H. Mueller and H. Navelet, Nucl. Phys. **B282**, 727 (1987).
- [75] C. Marquet and C. Royon, Nucl. Phys. **B739**, 131 (2006), arXiv:hep-ph/0510266.

- [76] O. Kepka, C. Royon, C. Marquet, and R. B. Peschanski, Phys. Lett. **B655**, 236 (2007), arXiv:hep-ph/0609299.
- [77] O. Kepka, C. Royon, C. Marquet, and R. B. Peschanski, Eur. Phys. J. **C55**, 259 (2008), arXiv:hep-ph/0612261.
- [78] C. Marquet and C. Royon (2007), arXiv:0704.3409 [hep-ph].
- [79] J. Bartels, D. Colferai, and G. P. Vacca, Eur. Phys. J. **C24**, 83 (2002), arXiv:hep-ph/0112283.
- [80] J. Bartels, D. Colferai, and G. P. Vacca, Eur. Phys. J. **C29**, 235 (2003), arXiv:hep-ph/0206290.
- [81] A. V. Kotikov and L. N. Lipatov, Nucl. Phys. **B582**, 19 (2000), arXiv:hep-ph/0004008.
- [82] G. P. Salam, JHEP **07**, 019 (1998), arXiv:hep-ph/9806482.
- [83] M. Ciafaloni, D. Colferai, G. P. Salam, and A. M. Stasto, Phys. Rev. **D68**, 114003 (2003), arXiv:hep-ph/0307188.
- [84] A. Sabio Vera, Nucl. Phys. **B722**, 65 (2005), arXiv:hep-ph/0505128.
- [85] D0 Collaboration, S. Abachi *et al.*, Phys. Rev. Lett. **77**, 595 (1996), arXiv:hep-ex/9603010.
- [86] V. Del Duca and C. R. Schmidt, Phys. Rev. **D49**, 4510 (1994), arXiv:hep-ph/9311290.
- [87] W. J. Stirling, Nucl. Phys. **B423**, 56 (1994), arXiv:hep-ph/9401266.
- [88] J. Bartels, V. Del Duca, and M. Wusthoff, Z. Phys. **C76**, 75 (1997), arXiv:hep-ph/9610450.
- [89] P. Aurenche, R. Basu, and M. Fontannaz (2008), arXiv:0807.2133 [hep-ph].
- [90] B. L. Combridge and C. J. Maxwell, Nucl. Phys. **B239**, 429 (1984).
- [91] L. N. Lipatov, Sov. J. Nucl. Phys. **23**, 338 (1976).
- [92] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977).
- [93] I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978).
- [94] J. G. Contreras, R. B. Peschanski, and C. Royon, Phys. Rev. **D62**, 034006 (2000), arXiv:hep-ph/0002057.
- [95] C. Marquet, R. B. Peschanski, and C. Royon, Phys. Lett. **B599**, 236 (2004), arXiv:hep-ph/0407011.

- [96] M. Ciafaloni, Phys. Lett. **B429**, 363 (1998), arXiv:hep-ph/9801322.
- [97] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C46**, 27 (2006), arXiv:hep-ex/0508055.
- [98] C. D. White, R. B. Peschanski, and R. S. Thorne, Phys. Lett. **B639**, 652 (2006), arXiv:hep-ph/0606169.
- [99] X. e. a. Aslanoglou, Eur. Phys. J. **C52**, 495 (2007), arXiv:0706.2641 [physics.ins-det].
- [100] L. Lonnblad, Comput. Phys. Commun. **71**, 15 (1992).
- [101] ZEUS Collaboration, S. e. a. Chekanov, Phys. Lett. **B632**, 13 (2006), arXiv:hep-ex/0502029.
- [102] X. Aslanoglou, A. Cyz, N. Davis, D. d'Enterria, E. Gladysz-Dziadus, C. Kalfas, Y. Musienko, A. Kuznetsov, and A. D. Panagiotou (2007), arXiv:0706.2576 [physics.ins-det].

# HERA Results

*Katerina Müller, Hanno Perrey, Thomas Schörner-Sadenius*

Jet production measurements at HERA allow detailed tests of our understanding of perturbative QCD. The concepts of factorization, of the perturbative expansion of the cross section and of PDF universality can all be tested. In addition, the strong coupling constant,  $\alpha_s$ , can be extracted from HERA jet production data. A further issue of particular relevance to the LHC, is the possible effects of multi-parton interactions and the underlying event. At HERA, such effects can be studied in resolved photoproduction events, since the hadronic structure of the photon means that photon-proton collisions are similar, in some respects, to hadron-hadron collisions.

The production of isolated photons, produced directly in the hard interaction, also provides a very sensitive probe of perturbative QCD, since the photons are largely insensitive to the effects of hadronisation. A good understanding of the Standard Model production mechanism of isolated photons is also important for searches of new particles decaying to photons at hadron colliders (eg.  $H \rightarrow \gamma\gamma$ ).

In this section, some recent HERA measurements of both jet production (mainly in the photoproduction regime), and of isolated photons, are reviewed. In addition, some recent measurements sensitive to the effects of multi-parton interactions and the underlying event are discussed.

## 1 HERA results on jets and prompt photons in photoproduction

*Authors: Hanno Perrey, Thomas Schörner-Sadenius*

In photoproduction at HERA, a quasi-real photon emitted from the incoming electron collides with a parton from the incoming proton. In such events, hadronic jets and also prompt (meaning: radiated by one of the outgoing quarks) photons can be produced. The photoproduction of hadronic jets can be classified into two types of processes in leading-order (LO) QCD: direct and resolved. In direct processes, the entire photon and its momentum participate in the hard scatter (left side of Fig. 1), while resolved processes involve a photon acting as a source of quarks and gluons, with only a photon momentum fraction  $x_\gamma$  participating in the hard scatter (right side of Fig. 1). It is due to the presence of resolved events that HERA data might be useful for further constraining the photon PDFs.

This contribution presents a review of some recent results on jet (and prompt photon) photoproduction at HERA. Some emphasis is placed on the prospects of using such measurements to constrain the proton and the photon PDFs. In addition, some results on multi-parton

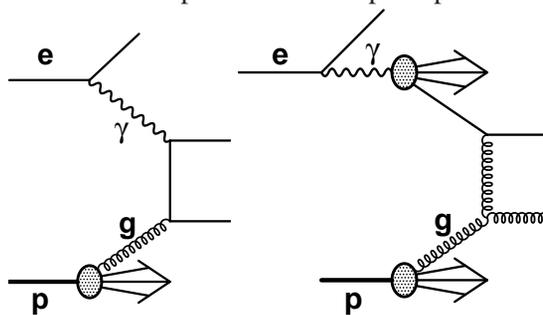


Fig. 1: Feynman diagrams of direct and resolved dijet photoproduction at leading order.

interactions and the underlying event are briefly summarised. It should be pointed out that most of these results use data only from the HERA-I data taking period, such that an improvement in statistical precision is to be expected by making use of all available data.

### 1.1 The concept of the resolved photon

Fig. 1 shows Feynman diagrams for direct (left) and resolved (right) photoproduction of dijets. Statistically, direct events are dominated by quark propagators whereas resolved events are mostly characterized by gluon propagators. This difference should lead to a distinctly different angular behaviour of the final-state jets: whereas the quark propagator (quarks being spin-1/2 particles) should lead to a distribution in the cosine of the centre-of-mass scattering angle,  $\cos \theta^*$ , like  $(1 - |\cos \theta^*|)^{-1}$ , in the gluon case a distribution like  $(1 - |\cos \theta^*|)^{-2}$  is expected. In other words, the cross section of the resolved part is expected to rise more rapidly towards higher  $\cos \theta^*$  than that of the direct part. Fig. 2 shows the experimental evidence [1]: the dijet cross section as a function of  $\cos \theta^*$  for a direct-enriched (left) and a resolved-enriched data sample (right). It is obvious that the above predictions are fulfilled, the resolved distribution rising much more rapidly than the direct one. These distributions thus form an important test of the concept of the resolved photon. Similar results have also been obtained by the ZEUS collaboration [2].

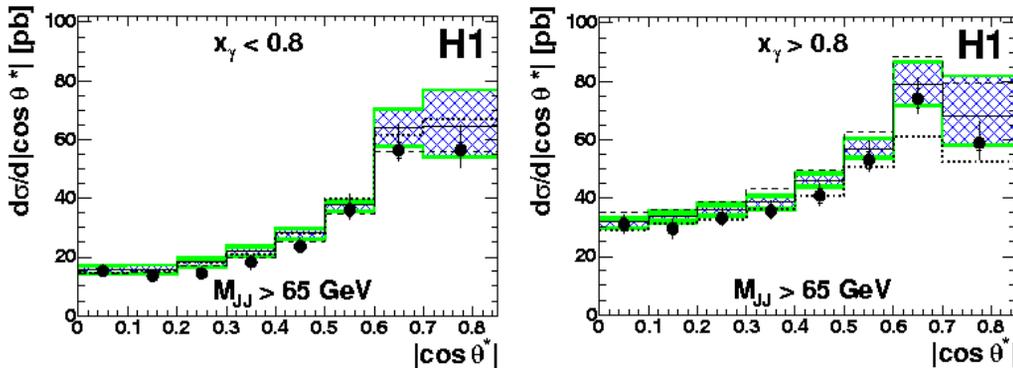


Fig. 2: Photoproduction dijet cross section as function of centre-of-mass scattering angle,  $\cos \theta^*$ , for direct- (left) and resolved-enriched (right) samples [1].

In the above discussion, the distinction between direct and resolved data samples has been made. On the theoretical side, this distinction is meaningful only at LO. On the experimental side, the distinguishing observable  $x_\gamma$  is not directly accessible but has to be reconstructed from the two final-state jets in much the same way as the proton's momentum fraction  $x_p$  entering the hard scattering,

$$x_\gamma = \frac{E_{T,1}e^{-\eta_1} + E_{T,2}e^{-\eta_2}}{2yE_e}, \quad x_p = \frac{E_{T,1}e^{+\eta_1} + E_{T,2}e^{+\eta_2}}{2E_p},$$

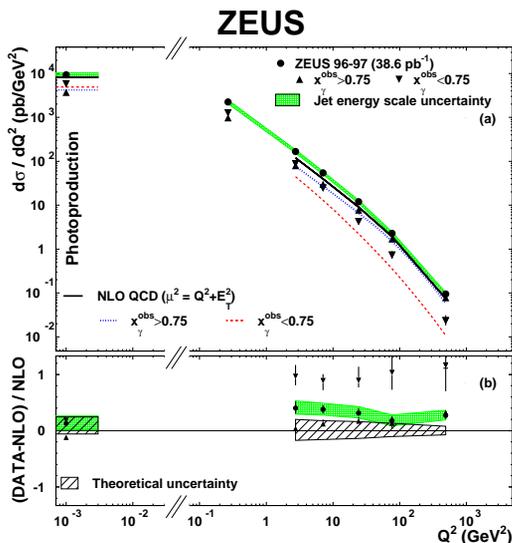


Fig. 3: Fraction of resolved events as a function of  $Q^2$  [3].

where  $E_{T,i}$  and  $\eta_i$  are the jet transverse energies and pseudorapidities,  $y$  is the inelasticity (characterizing the energy loss of the scattered electron) and  $E_e$  and  $E_p$  are the electron and proton beam energies. Typically, the resolved regime is defined to comprise values of  $x_\gamma$  between 0 and 0.75 or 0.8.

Note, however, that the phenomenon of the resolved photon is not strictly confined to the photoproduction regime. Also, the virtual photon entering into deep inelastic scattering events can exhibit a hadronic substructure, leading to a resolved contribution to DIS. The ZEUS collaboration has evaluated the fraction of resolved events in both photoproduction and DIS, measuring the fraction of dijet events with  $x_\gamma$  below and above 0.75. The results are shown in Fig. 3, as a function of the photon virtuality  $Q^2$  [3], and are compared to NLO QCD calculations. It is found that even at large  $Q^2$  values (highly virtual photons), there is a significant contribution from resolved events and that for DIS this component of the data is not correctly described by the QCD predictions (which do not include any resolved photon option). In contrast, the resolved contribution to photoproduction is well described by NLO QCD.

## 1.2 Jet cross sections in photoproduction

Numerous measurements of inclusive-jet, dijet and multijet cross sections have been performed by the HERA experiments. A very recent result [4] is presented in Fig. 4, which shows the dijet cross section as a function of the mean dijet transverse energy,  $\overline{E}_T$ , for a sample enhanced in direct and a sample enhanced in resolved events. The measurements are compared to an NLO QCD prediction using two different parameterisations of the photon PDFs.

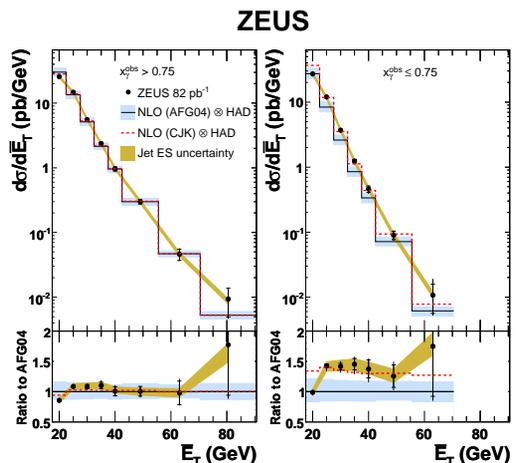


Fig. 4: Photoproduction dijet cross sections as functions of the mean transverse jet energy,  $\overline{E}_T$ , for a direct and a resolved enriched sample [4].

The data in the direct regime are especially well described by the theory (on the level of 10% or better), as can be seen in the bottom left part of the figure, which shows the ratio of data over NLO prediction. The uncertainties here are dominated by the theoretical uncertainty which is of the order of 15%. The situation in the resolved regime is slightly more complicated and will be discussed in some more detail below.

Many more examples of photoproduction jet cross sections and their successful description by NLO QCD exist. Fig. 6 [1] shows the dijet cross section as a function of  $x_p$  in different regions of  $x_\gamma$  and jet pseudorapidity. Both the momentum fractions and the pseudorapidity distributions of the jets are sensitive to the momentum distributions of partons inside the proton, making these measurements important tests of QCD. It can be seen again that the data are very well described by NLO QCD, on the level of 10%, which are well covered by the combined uncertainties. Only for large  $x_p$  values, with both jets going forward ( $\eta_{1,2} > 1$ ), do some deviations between data and theory occur (see Fig. 6, bottom right). These differences might be explained by the large uncertainties on the proton PDFs for large momentum fractions.

Fig. 5 [4] shows, for the same data sample as in Fig. 4, the cross section as a function of the mean jet pseudorapidity,  $\bar{\eta}$ . The data are again shown separately for direct- and resolved-enhanced samples and are compared to NLO QCD predictions using different parameterisations of the photon PDFs. For the direct case, the description of the data by the theory is again excellent.

The demonstrated good performance of NLO QCD in describing photoproduction data (especially in the direct regime) gives confidence in the theory, thus rendering possible the extraction of QCD parameters like the strong coupling constant,  $\alpha_s$ , or the proton and photon PDFs from the data. One example of the former is given in [5], where a value  $\alpha_s(M_Z) = 0.1224 \pm 0.0001(\text{stat.})_{-0.0019}^{+0.0022}(\text{exp.}) \pm_{-0.0042}^{+0.0054}(\text{th.})$  was extracted from an inclusive-jet measurement in photoproduction. Many other precision extractions of  $\alpha_s$ , from jet measurements in both DIS and photoproduction, also exist. The impact of jet cross sections in constraining the PDFs is discussed in the next section.

Recently, the symmetry group underlying the strong interaction has been studied by measuring angular correlations in direct-enriched three-jet photoproduction events [6]. Fig. 7 shows the differential cross sections as functions of the angular correlations between the three jets in the final state and the proton-beam direction. The data are compared to LO calculations based on different symmetry groups, illustrating the sensitivity of the data to the colour configuration. While the measured angular correlations are consistent with the prediction of SU(3), they disfavour other symmetry groups such as SU( $N$ ) in the limit of large  $N$ . The differences between SU(3) and U(1)<sup>3</sup> were found to be smaller than the current statistical uncertainties.

### 1.3 Jets in Photoproduction and the Proton and Photon PDFs

In [4], both the theoretical uncertainties on dijet cross sections and their sensitivity to the gluon density in the photon and the proton have been investigated in great detail. As is highlighted in Fig. 8 (left), for a special choice of kinematics, there are regions in which the proton PDF uncertainty (indicated as the region between the two solid lines) is as large as, or even larger than, the combined theoretical uncertainties due to the choice of renormalization scale, factorization scale,

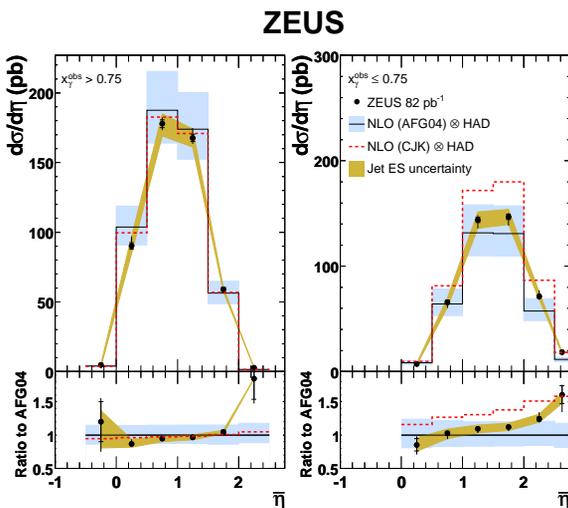


Fig. 5: Photoproduction dijet cross sections as functions of the mean jet pseudorapidity,  $\bar{\eta}$ , for a direct- and a resolved enriched sample [4].

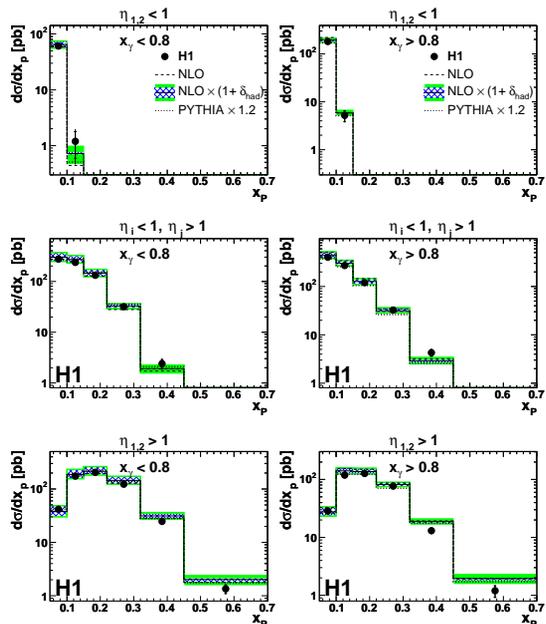


Fig. 6: Photoproduction dijet cross sections as functions of the momentum fraction  $x_p$  for a direct- and a resolved enriched sample [1].

$\alpha_s$ , and hadronisation correction. Also, the uncertainty due to the very imprecise knowledge of the photon PDF may be very large, reaching values of up to 60%, as indicated by the dashed line in the figure, which shows the difference in the cross section prediction between two different photon PDF parameterisations. Dijet data do, therefore, have the potential to further constrain both quark and gluon densities in the proton and the photon. This is indicated in Fig. 8 (right), which shows the fraction of gluon-induced events on the proton side (dark dashed line) and on the photon side according to two different photon PDF parameterisations (light dashed and solid lines). The amount of gluon-induced events can be as large as 60%, depending on the detailed kinematics under consideration.

The large discrepancies between different photon PDF parameterisations are also visible in the comparison of NLO QCD predictions with dijet cross sections in the resolved regime, like in Fig. 4 (right) or Fig. 5 (right). For example, in Fig. 4, the resolved dijet cross section can be approximately described by the NLO prediction using the CJK photon PDF parametrization, but not by the AFG04 parametrization, which is off by up to 40–50%. This difference highlights the potential of the data to further constrain the photon PDFs.

A first example of the benefit of jet photoproduction data on determinations of the proton PDFs is given in Figs. 9 and 10. Fig. 9 shows, for an older measurement of photoproduction dijet cross sections [2], the ratio of the measured cross sections over the NLO prediction. An

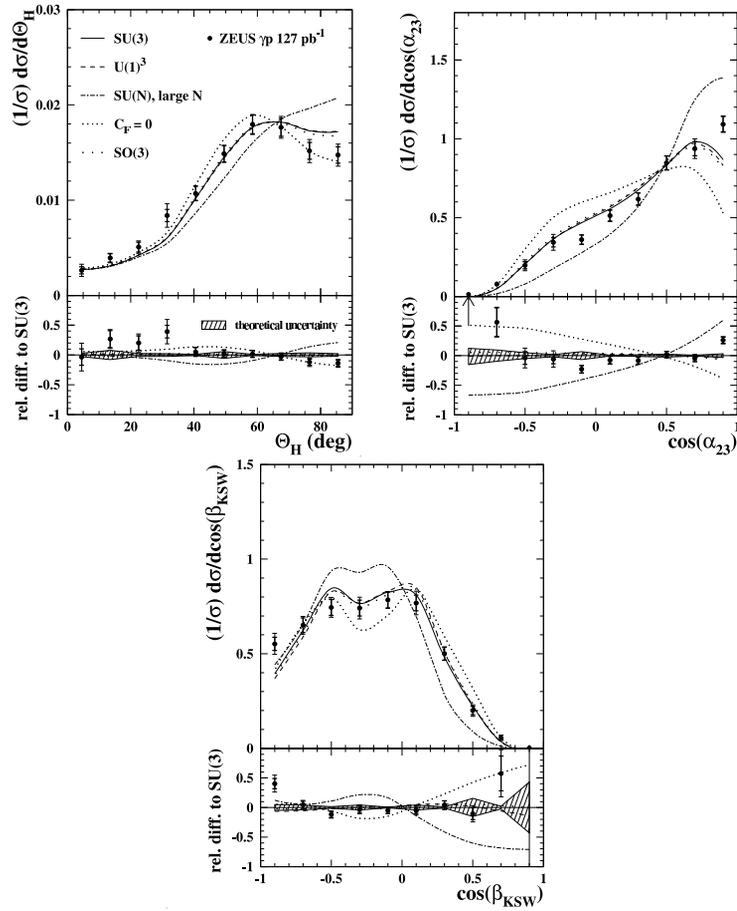


Fig. 7: Photoproduction trijet cross sections as functions of angular correlation between the jets and the proton-beam direction, for a direct enriched sample [6].

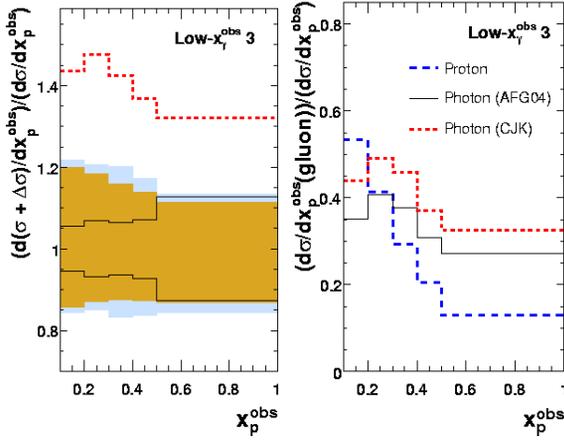


Fig. 8: Left: Theoretical uncertainties on photoproduction dijet cross sections in a special kinematic region [4]. The uncertainties are the total (outer shaded band), that from varying  $\mu_R$  (inner shaded band), the proton PDF uncertainties from the ZEUS-JETS fit (solid lines) and the difference from using the CJK photon PDF, rather than AFG04 (dashed line). Right: Gluon-induced contributions to photoproduction dijet cross sections in the same kinematic region.

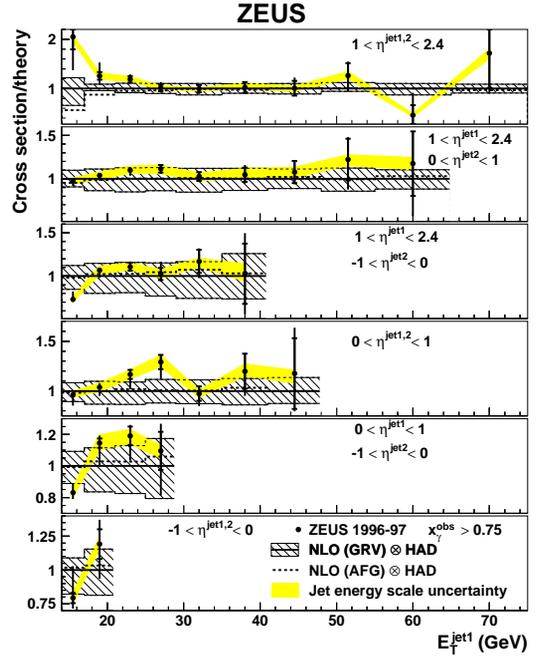


Fig. 9: Ratio of data over NLO QCD predictions for photoproduction dijet events in the direct regime as function of jet transverse energy and pseudorapidity,  $E_T$  and  $\eta$  [2].

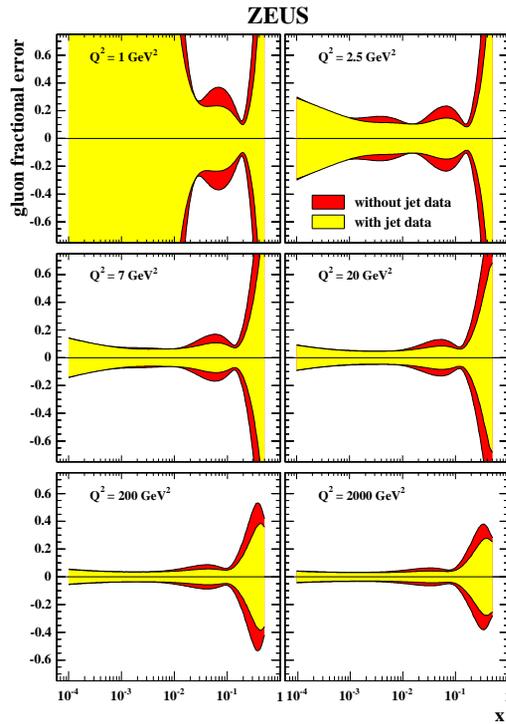


Fig. 10: Improvement of the gluon density by using jet data in NLO QCD fits of the PDFs [7].

overall good description is found, with data and theory in agreement almost everywhere within the combined theoretical and experimental uncertainties. These data (together with data from DIS jet analyses) have been used as additional inputs (besides the usual inclusive  $F_2$  data) to an NLO QCD PDF fit [7]. The success of this fit is demonstrated in Fig. 10, which shows the fractional gluon density uncertainty as a function of the proton momentum fraction  $x$  in different regions of  $Q^2$ . The uncertainty without the use of jet data is given by the dark shaded area, and the result including jet data is given by the light shaded area. An improvement in the uncertainty of up to 35% is clearly visible, especially in the region of medium-to-high- $x$  values.

The aim now is to further improve the proton PDFs (and here especially the gluon density at high values of  $x$ , since this is particularly important for LHC physics) by using more precise measurements or cross sections measured in different kinematic regions, from both photoproduction and DIS. Constraining the photon PDFs will be technically even more demanding, partly because of lack of a consistent PDF error treatment for the photon PDF, and partly because of the increased experimental and theoretical uncertainties for the resolved regime.

#### 1.4 Photoproduction of isolated photons

The production of isolated, or “prompt”, photons from the hadronic final state offers an alternative access to the QCD dynamics in  $ep$  scattering, with different systematic uncertainties and reduced effects from hadronisation. Isolated photon production has been measured by ZEUS and H1 in both photoproduction [8–10] and DIS [11, 12]. In addition to inclusive measurements of isolated photons, cross sections of photons in association with jets are also often measured, for which there are currently NLO QCD predictions available. In Sec. 2, a detailed description of the latest HERA prompt photon measurement [12] in DIS is presented. Here, only a very brief summary of a recent measurement in photoproduction is given [10].

Fig. 11 [10] shows the photon plus jet cross section, in photoproduction, as a function of  $E_T^\gamma$  and  $\eta^\gamma$ . The data are compared to NLO QCD calculations (shaded band) and to calculations based on the  $k_T$ -factorization approach and unintegrated parton densities (hatched band). The results demonstrate that the  $k_T$ -factorization approach gives the closest agreement with the data, especially in the low  $E_T^\gamma$  and forward  $\eta^\gamma$  regions. The fact that NLO QCD does not describe the data at low values of  $E_T^\gamma$  (or  $E_T^{jet}$ ) currently precludes the use of the low-transverse-energy data in constraining the PDFs or  $\alpha_s$ .

#### 1.5 The underlying event and multi-parton interactions

Resolved photon-proton interactions may, in some respects, be regarded as hadron-hadron collisions, with all the additional features with respect to direct interactions. In particular, it is possible in hadron-hadron collisions to have multiple interactions of pairs of partons (so-called ‘multi-parton interactions’ or ‘MPI’), which may populate the hadronic final state with additional soft or hard jets, or additional energy flow throughout the detector. This effect may alter the final state significantly, making it necessary to model it adequately in the Monte Carlo programs used in the analyses. There exist various MPI model implementations in standard generators, such as HERWIG and PYTHIA, which can be tested against data, or whose parameters can be adjusted to describe the data. Here, some recent examples of multi-parton interaction studies at HERA,

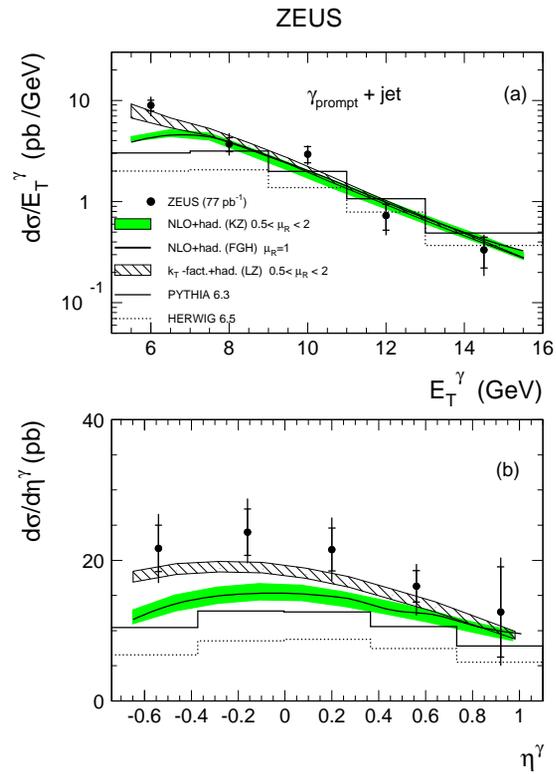


Fig. 11: The  $\gamma + \text{jet}$  differential cross sections in photoproduction as functions of  $E_T^\gamma$  and  $\eta^\gamma$  compared to theoretical QCD calculations [10].

from H1 and ZEUS, are briefly reviewed. Note that a more detailed discussion of the H1 results can be found in the WG5 section of these proceedings.

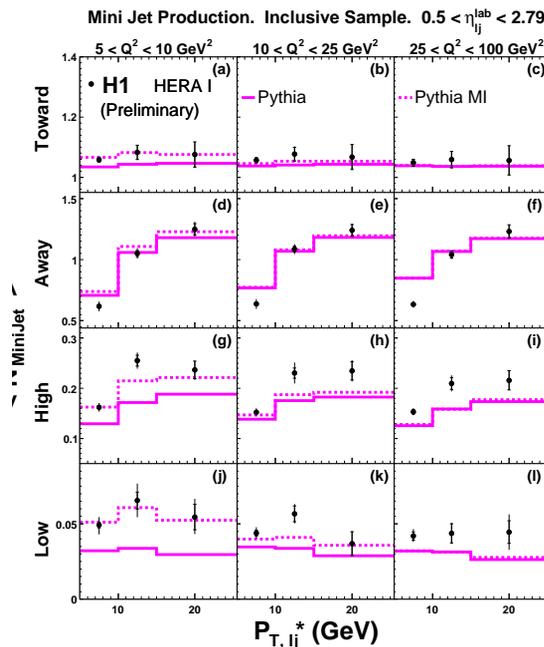


Fig. 12: Mean number of minijets as function of the leading jet  $p_T$  [13].

Fig. 12 shows the recent H1 measurement [13] of the mean number of ‘minijets’ (i.e. soft jets with transverse energies above a very low cut of 3 GeV) in DIS events, with at least one hard jet. The data are shown as a function of this leading jet’s transverse momentum,  $p_{T,lj}^*$ , in different regions of  $Q^2$ , and different regions of azimuthal angle with respect to the leading jet’s azimuth. The so-called ‘Towards’ and ‘Away’ regions are expected to be mostly populated by the results of the first and hardest parton-parton scattering in the event, and the dijet system coming from this scattering should be separated by an azimuthal angle of about  $\pi$ . In contrast, the ‘High’ and (especially) the ‘Low’ regions<sup>1</sup>, should be particularly sensitive to MPI effects which, in these regions, are not masked out by the harder energy depositions from the leading jet pair. It can be observed that the PYTHIA model with MPI effects switched on (‘PYTHIA MI’) is in rather good agreement with the data in almost all regions. In contrast, PYTHIA without MPI modelling (‘PYTHIA’) fails to describe the data in the low  $Q^2$  ‘High’ and ‘Low’ regions, consistent with the hypothesis of MPI effects dominating in these regions of phase space. Overall, the results indicate that the data strongly favour the models including MPI effects.

Another recent H1 measurement [14], this time of dijets in photoproduction, leads to similar conclusions. Fig. 13 shows the charged particle multiplicity for a resolved-enriched data sample in the ‘High’ (left) and ‘Low’ (right) regions as function of the transverse momentum

<sup>1</sup>Note that ‘High’ and ‘Low’ refer to the amount of deposited energy in the two regions.

of the leading jet. Again, PYTHIA without MPI modelling (‘PYTHIA NMI’) does not give a sufficient description of the data while PYTHIA including MPI (‘PYTHIA MI’) models the data well.

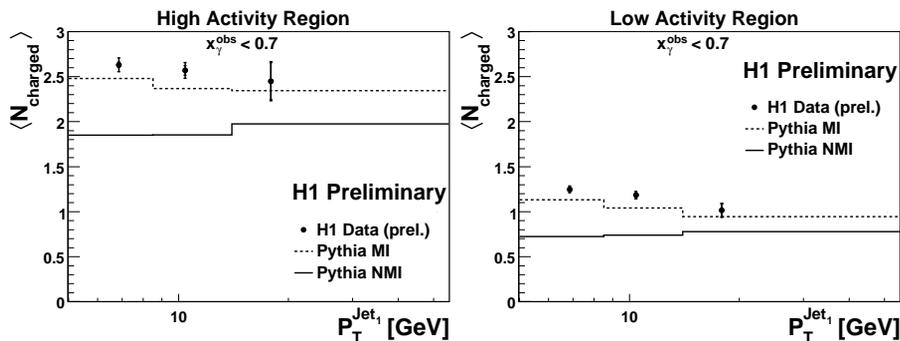


Fig. 13: Charged particle multiplicity for resolved-enriched photoproduction dijets as function of the transverse momentum of the first jet [14].

A similar statement can also be derived from a recent measurement of three- and four-jet photoproduction [15] from ZEUS. Fig. 14 shows the four-jet photoproduction cross section as a function of  $x_\gamma$ , and compares it to the predictions of HERWIG and PYTHIA, with and without the inclusion of MPI modelling. The results indicate that the two predictions including MPI effects (‘HERWIG+MPI’ and ‘PYTHIA+MPI’) are able to adequately reproduce the data, whereas the models without MPI grossly underestimate the measured cross section in the resolved-enhanced regime (i.e.  $x_\gamma < 0.8$ ). The effect is particularly drastic for low energy scales and low multijet invariant masses, a fact which is illustrated in Fig. 14, where the four-jet mass was required to be between 25 and 50 GeV. It should also be noted that the parameters of the “HERWIG+MPI” model have been tuned to these multijet data, while the “PYTHIA+MPI” parameters have not.

While the HERA measurements clearly favour models in which MPI are included, the data so far do not have the power to specify more precisely the mechanism underlying MPI effects, or to shed light on the energy evolution of MPI effects when going (for example) from TEVATRON to LHC centre-of-mass energies. The models in use so far are rather crude and will have to be replaced by more realistic models and calculations which take correctly into account features like multi-parton exchanges between photon and proton, correlations between these exchanges, etc.

## 2 Measurement of Isolated Photon Production in Deep Inelastic Scattering at HERA

*Author: Katharina Müller*

Both the H1 and ZEUS Collaborations have previously measured [8–10] prompt photon cross sections in photoproduction (see, for example, Sec. 1.4). An analysis of the isolated photon cross section in deep inelastic scattering (DIS), with  $Q^2 > 35 \text{ GeV}^2$ , has also been published by ZEUS [11]. The present contribution describes the most recent HERA measurement of isolated photon production in DIS ( $e + p \rightarrow e + \gamma + X$ ) from H1.

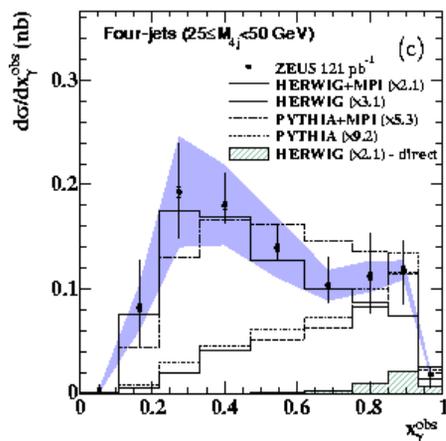


Fig. 14: Photoproduction four-jet cross section as function of  $x_\gamma$  [15].

The measurement of isolated photons in deep inelastic scattering provides a test of QCD in a kinematic range with two hard scales, the four-momentum transfer squared,  $Q^2$ , of the exchanged virtual photon and the transverse energy of the emitted photon. In DIS, the final state photon is emitted by a quark ( $QQ$  subprocess), by wide angle radiation from the lepton ( $LL$  subprocess) or by interference ( $LQ$  subprocess). The  $QQ$  contribution is dominated by the direct radiation of the photon from the quark involved in the parton level process, but also contains the contribution from quark fragmentation to a photon [16, 17]. Since the photon and the scattered electron are well separated in the present analysis, low angle QED radiation is suppressed. The  $LL$  and the more interesting  $QQ$  process can not be distinguished experimentally on event basis, but only by comparison with predictions.

The measurement of isolated photons in DIS presented here is based on a total integrated luminosity of  $227 \text{ pb}^{-1}$ . Full details of the analysis can be found in [12]. Photon candidates with transverse energy  $3 < E_T^\gamma < 10 \text{ GeV}$  and pseudorapidity  $-1.2 < \eta^\gamma < 1.8$  are selected in DIS events in the kinematic regime  $4 < Q^2 < 150 \text{ GeV}^2$ , inelasticity  $y > 0.05$  and a mass of the hadronic system  $W_X > 50 \text{ GeV}$ . The cut on  $W_X$  removes events from elastic compton scattering. The photon candidates are then used together with the other particles in the event, with the exception of the scattered electron, to reconstruct jets using the  $k_T$  algorithm [18]. The isolation of the photon is ensured by requiring that it carries at least 90% of the transverse momentum of the jet containing the photon. The isolation criteria removes a large part of the background from decay products of neutral hadrons.

Photons are separated from the remaining neutral hadrons and their decay products by a multivariate analysis of the shapes of the calorimeter energy deposits.

Jet production in events with isolated photons and no additional jet or with at least one jet is also investigated. Hadronic jets are reconstructed for  $P_T^{jet} > 2.5 \text{ GeV}$ , the pseudorapidity range is restricted to  $-1.0 < \eta^{jet} < 2.1$ . All results are compared to a leading order (LO),  $\mathcal{O}(\alpha^3 \alpha_s^0)$ , calculation [19, 20]. The cross sections for a photon plus at least one jet are further

compared to a NLO,  $\mathcal{O}(\alpha^3\alpha_s^1)$ , calculation [21].

Differential cross sections  $d\sigma/d\eta^\gamma$  and  $d\sigma/dQ^2$  for the inclusive isolated photon cross section are shown in Fig. 15. Fig. 16 shows the differential cross sections  $d\sigma/d\eta^\gamma$  for isolated photons with no additional hadronic jet (a) and with hadronic jets (b). The uncertainty on the shower description gives the dominant contribution to the systematical error. The cross sections are shown together with the predictions by the LO calculation. A comparison to a NLO calculation is possible for the photon plus jets cross section. The calculations are corrected to hadron level. The corrections amount to, at most,  $-30\%$ .

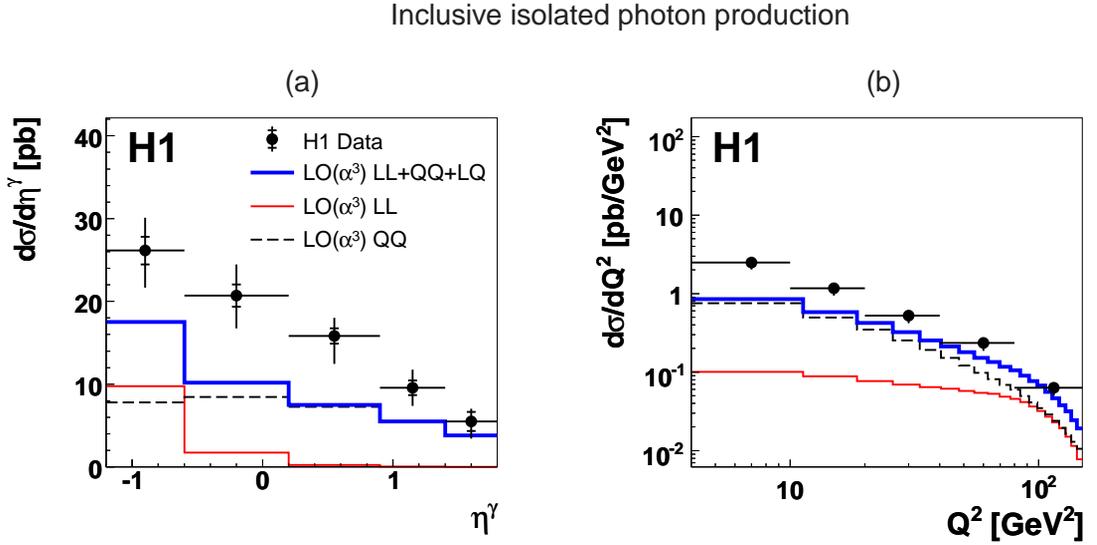


Fig. 15: Inclusive differential isolated photon cross sections  $d\sigma/d\eta^\gamma$  (a),  $d\sigma/dQ^2$  (b) for  $3 < E_T^\gamma < 10$  GeV,  $-1.2 < \eta^\gamma < 1.8$ ,  $W_X^2 > 2500$  GeV<sup>2</sup> and  $4 < Q^2 < 150$  GeV<sup>2</sup>. The inner error bars on the data points indicate the statistical error, the full error bars contain in addition the systematic errors added in quadrature. The cross sections are shown together with a leading order,  $\alpha^3\alpha_s^0$ , calculation [20] corrected for hadronisation effects, *LL* corresponding to radiation from the electron and *QQ* to radiation from the quark.

The LO  $\mathcal{O}(\alpha^3\alpha_s^0)$  calculation underestimates the inclusive cross sections by roughly a factor of two, most significantly at low  $Q^2$ . The relative contribution of radiation from the electron (*LL*) and the quark(*QQ*) depends strongly on  $\eta^\gamma$  and  $Q^2$ . At high and medium  $\eta^\gamma$  and low  $Q^2$ , radiation by the quark dominates. The shapes of the  $d\sigma/dE_T^\gamma$  and  $d\sigma/d\eta^\gamma$  distributions are described reasonably well. The comparison of data to the LO calculation in bins of  $\eta^\gamma$  show that the difference in normalisation can mainly be attributed to an underestimation of the *QQ* contribution [12].

The LO prediction also underestimates the production of isolated photons plus no hadronic jet and photons plus jets by a similar factor as for the inclusive measurement. The NLO  $\mathcal{O}(\alpha^3\alpha_s^1)$

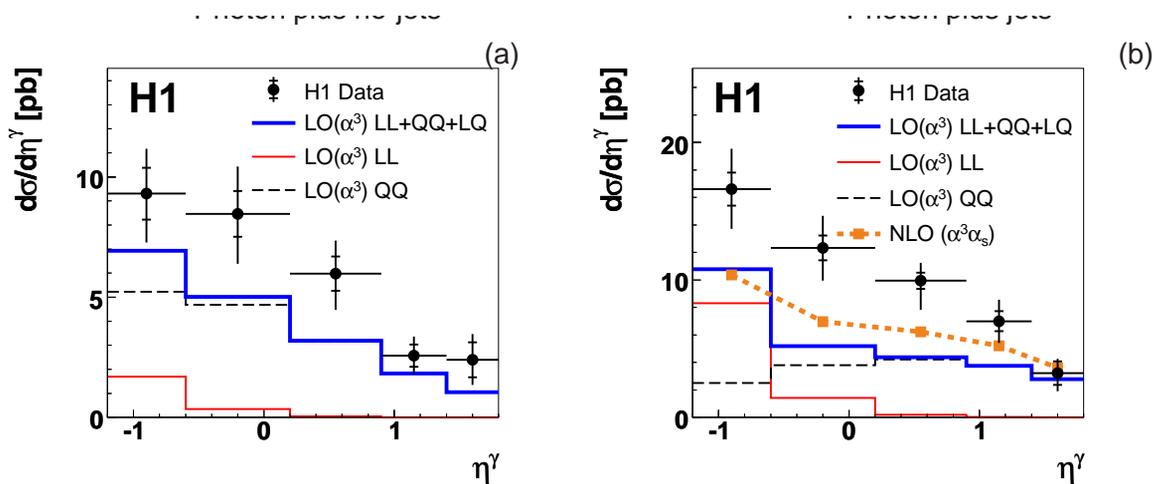


Fig. 16: Differential cross sections  $d\sigma/d\eta^\gamma$  for photon plus no-jets (a) and photon plus jets (b), with  $P_T^{jet} > 2.5$  GeV and  $-1.0 < \eta^{jet} < 2.1$ . The cross sections are compared to a leading order,  $\mathcal{O}(\alpha^3\alpha_s^0)$ , calculation [19] as in Figure 1. The photon plus jet sample is additionally compared to a NLO ( $\alpha^3\alpha_s$ ) calculation [21]. The bin averaged NLO cross sections are indicated by the squares.

prediction for photon plus jet is higher than the LO prediction, most significantly at low  $Q^2$ , but still underestimates the data. The NLO calculation describes the shapes of the differential cross sections reasonably well.

The  $LL$  contribution is largely suppressed for the sample with no additional jet due to the cut on  $W_X$ . The cross section for photon plus jet production is roughly two times higher than for photon plus no additional jet. This is in contrast to the inclusive  $ep \rightarrow eX$  cross section, where topologies with an additional jet are suppressed by  $\mathcal{O}(\alpha_s)$ . The similar cross sections for photon events with or without additional jets may be explained by the fact that both topologies correspond to the same order  $\mathcal{O}(\alpha_s^0)$  in perturbative QCD. [20] The cross section for photon plus jet production is roughly two times higher than for photon plus no-jets.

At leading order  $\mathcal{O}(\alpha_s^0)$ , the quark fragmentation contribution to the cross section enters only the sample with a photon and no hadronic jet [20]. Hence, the observed excess of the data can not solely be attributed to an underestimation of that contribution.

Further theoretical investigations are needed to understand the observed discrepancy between the measurements and the predictions, including for instance the calculation of higher order processes.

## References

- [1] H1 Collaboration, A. Aktas *et al.*, Phys. Lett. **B639**, 21 (2006), arXiv:hep-ex/0603014.

- [2] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C23**, 615 (2002), arXiv:hep-ex/0112029.
- [3] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C35**, 487 (2004), arXiv:hep-ex/0404033.
- [4] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Rev. **D76**, 072011 (2007), arXiv:0706.3809 [hep-ex].
- [5] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Lett. **B560**, 7 (2003), arXiv:hep-ex/0212064.
- [6] ZEUS Collaboration, . S. Chekanov (2008), arXiv:0808.3783 [hep-ex].
- [7] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C42**, 1 (2005), arXiv:hep-ph/0503274.
- [8] ZEUS Collaboration, J. Breitweg *et al.*, Phys. Lett. **B472**, 175 (2000), arXiv:hep-ex/9910045.
- [9] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C38**, 437 (2005), arXiv:hep-ex/0407018.
- [10] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C49**, 511 (2007), arXiv:hep-ex/0608028.
- [11] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Lett. **B595**, 86 (2004), arXiv:hep-ex/0402019.
- [12] H1 Collaboration, F. D. Aaron *et al.*, Eur. Phys. J. **C54**, 371 (2008), arXiv:0711.4578 [hep-ex].
- [13] H. Collaboration, *Minijet Production in Deep Inelastic Scattering at HERA* (unpublished). H1-PRELIM-07-032.
- [14] H. Collaboration, *Study of multiple interactions in photoproduction at HERA* (unpublished). H1-PRELIM-08-036.
- [15] ZEUS Collaboration, S. Chekanov *et al.*, Nucl. Phys. **B792**, 1 (2008), arXiv:0707.3749 [hep-ex].
- [16] K. K. *et al.*, Z. Phys. C **6**, 131 (1980).
- [17] M. Gluck, L. E. Gordon, E. Reya, and W. Vogelsang, Phys. Rev. Lett. **73**, 388 (1994).
- [18] S. D. Ellis and D. E. Soper, Phys. Rev. **D48**, 3160 (1993), arXiv:hep-ph/9305266.
- [19] A. Gehrmann-De Ridder, T. Gehrmann, and E. Poulsen, Phys. Rev. Lett. **96**, 132002 (2006), arXiv:hep-ph/0601073.

- [20] A. Gehrmann-De Ridder, T. Gehrmann, and E. Poulsen, Eur. Phys. J. **C47**, 395 (2006), arXiv:hep-ph/0604030.
- [21] A. Gehrmann-De Ridder, G. Kramer, and H. Spiesberger, Nucl. Phys. **B578**, 326 (2000), arXiv:hep-ph/0003082.

# Interactions at high gluon densities

Mark Strikman, Igor M. Dremin

In the previous section we mentioned the subject of gluon saturation. In this section we continue the discussion of effects due to high gluon densities. First we look at what HERA can teach us about the density of gluons in the impact parameter plane and how this will affect our understanding of processes in  $pp$  collisions at LHC. Then, in Sec. 2 we go on to heavy ion collisions and discuss effects of a dense gluon medium there, concentrating on the description of Cherenkov gluons.

## 1 HERA constrains for LHC MC generators and probing high gluon densities in $pp$ collisions using forward triggers

Author: Mark Strikman

In the high energy collisions the finite  $x$  component of the wave functions of the colliding hadrons is nearly frozen in transverse plane during the interaction process. Properties of produced final state depend strongly on whether hadrons collided at large impact parameter,  $b$  or head on. In particular for small  $b$  a chance for a parton to pass through high gluon density fields at a distance  $\rho$  from the center of the second nucleon (Fig. 1) is enhanced. The probability of multiple collisions parton collisions is enhanced as well.

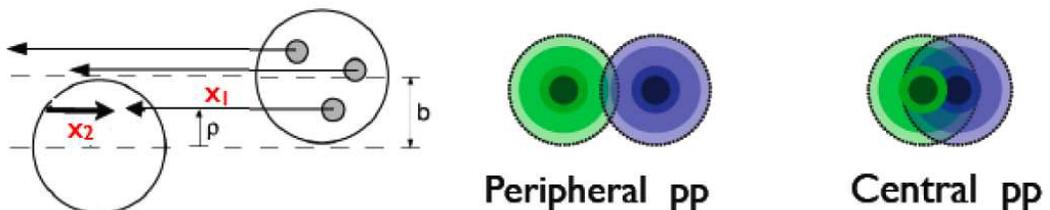


Fig. 1: Side and transverse views of  $pp$  collision.

The strength of the encountered gluon fields depends strongly on  $x$  of the parton - a parton with a given  $x_1$  and resolution  $p_t$  is sensitive to the partons in the target with  $x \geq x_2 = 4p_t^2/s_{NN}x_1$ . For fixed  $x_1$  characteristic  $x_2$  decrease  $\propto 1/s$ . For example at the LHC a parton with  $x_1 = 0.1, p_t = 2\text{GeV}/c$  resolves  $x > 10^{-6}$  while at the GZK energies such parton resolves  $x > 10^{-9}$  corresponding to huge gluon densities since a change of  $x$  by a factor of ten leads to an increase of gluon density by at least a factor of two.

Studies at HERA provided several important inputs which we discuss below: (i) transverse distribution of gluons in the nucleon, (ii) fluctuations of the strength of the gluon field in the nucleon, (iii) proximity to the black disk regime. When combined with information from the Tevatron collider they indicate also correlations of partons in the transverse plane.

These observations have a number of implications for the dynamics of  $pp$  collisions at LHC energies, which are most pronounced in the forward region. Hence we also discuss how to

trigger on central  $pp$  collisions and how to use such collisions for study of the small  $x$  dynamics at very small  $x$ .

### 1.1 Exclusive hard diffraction at HERA - implications for MC at the LHC

The QCD factorization theorem [1, 2] allow to determine the generalized gluon distribution in nucleon for small  $x$  from the DIS exclusive meson production at small  $x$  as well as from the production of onium states. The  $t$ -dependence of these distributions is connected via Fourier transform to the transverse distribution of gluons in a nucleon for a given  $x$ . The data confirm our prediction of convergence of the  $t$ -slopes for different mesons with increase of  $Q^2$  and weak dependence of the  $t$ -slope for the  $J/\psi$ -meson production on  $Q^2$ . Accordingly, this allows to determine the transverse distribution of gluons as a function of  $x$  (for review and references see [3]). It can be approximated as

$$F_g(x, \rho) = \frac{m_g^2}{2\pi} \left( \frac{m_g(x)\rho}{2} \right) K_1(m_g(x)\rho), \quad (1)$$

where  $K_1$  denotes the modified Bessel function. We find  $m_g^2(x = 0.05) \sim 1.1 GeV^2$  which corresponds to a much more narrow transverse distribution than given by the electro-magnetic form factors. The radius of the gluon distribution grows with decrease of  $x$  reaching the value comparable to the e.m. radius for  $x \sim 10^{-4}$  ( $m_g^2 \sim 0.7 GeV^2$ ).

Hence analysis of the HERA data suggests that the transverse gluon distribution,  $F_g(x, \rho)$ , significantly broadens with decrease of  $x$ . At the same time the current MC models of  $pp$  collisions assume that transverse parton distributions do not depend on  $x$ . Also, in the PYTHIA MC [4] it is assumed that two transverse scales are present in the  $\rho$ -dependence of  $F_g$ . It is not clear whether this assumption is consistent with Eq. (1) and correspondingly with the data on the exclusive  $J/\psi$  production.

Knowledge of  $F_g(x, \rho)$  allows to calculate the rate of the production of four jets due to double parton collisions in the  $pp$  scattering assuming that the double parton distribution is given by a product of single parton distributions. Using Eq. (1) we find the rate which is a factor of two smaller than observed in the Tevatron experiment [5, 6]. This implies presence of the transverse correlations between partons.

One of the sources of fluctuations is fluctuations of the overall size of the initial parton configurations. In the high energy scattering different initial configurations in the colliding nucleons can be considered as frozen. Studies of the soft inelastic diffraction indicate that the strength of the interaction for different configurations in nucleons fluctuates rather strongly. Presumably significant contribution to these fluctuations comes from the fluctuation of the size of these configurations. One also expects that parton distributions in different configurations should differ as well.

In ref. [7] we deduced the model-independent relation which allows one to infer the small  $x$  fluctuations of the gluon density from the observable ratio of inelastic ( $\gamma *_{L} + p \rightarrow VM + X$ ) and elastic ( $\gamma *_{L} + p \rightarrow VM + p$ ) diffractive vector meson production at  $t = 0$ :

$$\omega_g \equiv \frac{\langle G^2 \rangle - \langle G \rangle^2}{\langle G \rangle^2} = \left[ \frac{d\sigma_{\text{inel}}}{dt} / \frac{d\sigma_{\text{el}}}{dt} \right]_{t=0}^{\gamma_L^* p \rightarrow VX}. \quad (2)$$

So far there have been no dedicated experimental studies of this ratio. Overall data suggest that  $\omega_g \sim 0.2$  for  $Q^2$  of few  $\text{GeV}^2$  and  $x \leq 10^{-3}$  which corresponds to rather large fluctuations of the gluon density. We also proposed a simple model based on information on the fluctuations of the strength of the strong interaction which allows to reproduce the magnitude of  $\omega_g$ .

Correlations between fluctuations of the parton densities and the soft–interaction strength have numerous potential implications for high–energy  $pp/\bar{p}p$  collisions with hard processes. One example is the relative probability of double binary parton–parton collisions.

The QCD evolution leads to a drop of the fluctuations with an increase of virtuality. As a result in the case of double scattering configurations, the main effect for the overall rate is due to fluctuations of the size of the transverse area of the configurations. The contribution of configurations of size smaller than average is enhanced leading to a rather modest enhancement of the rate of four jet production  $\sim 10\text{--}15\%$ , which accounts for a small fraction of the remaining discrepancy with the CDF value<sup>1</sup>. However the size of configurations involved in the multijet double / triple scattering trigger is much smaller than the average size, leading to modification of the hadron product in the fragmentation region, long range fluctuations of multiplicity, etc.

Small effect from global fluctuations indicates that other dynamical mechanisms must be responsible for the enhancement of multi–parton collisions, *e.g.* local transverse correlations between partons as suggested by a “constituent quark” picture of the nucleon [3].

## 1.2 Onset of the black regime in the interaction of fast partons

Interactions of virtual photons with nucleons at HERA can be represented as superposition of the interaction of  $q\bar{q}$  dipoles of sizes given by the square of the corresponding photon wave function. The cross section of the inelastic interaction of a  $q\bar{q}$  or gluon dipole can be written as

$$\sigma^{\text{q}\bar{\text{q}}\text{--hadron}}(x, d^2) = \frac{\pi^2}{4} F^2 d^2 \alpha_s(Q_{\text{eff}}^2) x G_T(x, Q_{\text{eff}}^2). \quad (3)$$

Here  $F^2 = 4/3$  is the Casimir operator of the fundamental representation of the  $SU(3)$  gauge group. Furthermore,  $\alpha_s(Q_{\text{eff}}^2)$  is the LO running coupling constant and  $G_T(x, Q_{\text{eff}}^2)$  the LO gluon density in the target. They are evaluated at a scale  $Q_{\text{eff}}^2 \approx \lambda d^{-2}$ , where  $\lambda = 5 \div 9$  can be determined from NLO calculations or from phenomenological considerations.

Since the gluon density rapidly increases with decrease of  $x$  while the transverse radius of the nucleon grows rather slowly, one expects based on Eq. (3) that interaction should approach the black disk regime of complete absorption at sufficiently large energies. To determine the proximity to this limit it is convenient to study the amplitude of the dipole - nucleon scattering,  $A^{dp}(s, t)$  which can be inferred from analysis of the data on the total DIS cross section and data on exclusive production of vector mesons [8].

Introducing impact parameter representation of the amplitude

$$A^{dp}(s, t) = \frac{is}{4\pi} \int d^2b e^{-i(\Delta_{\perp} \mathbf{b})} \Gamma^{dp}(s, b) \quad (t = -\Delta_{\perp}^2), \quad (4)$$

---

<sup>1</sup>Note that the CDF measurements correspond to relatively large  $x$  where the “radiative” model of the gluon density fluctuations we developed may not be applicable and where no data on the hard inelastic exclusive diffraction are available. However, if the gluon strength is larger for configurations of larger size, it would lead to reduction of already rather small enhancement of the rate of multiple collisions.

we can determine  $\Gamma^{dp}(s, b)$  which is referred to as the profile function. In the situation when elastic scattering is the “shadow” of inelastic scattering, the profile function at a given impact parameter is restricted to  $|\Gamma^{dp}(s, b)| \leq 1$ . The probability of the inelastic interaction for given  $b$

$$P_{inel}(b) = 1 - \left| 1 - \Gamma^{dp}(s, b) \right|^2, \quad (5)$$

is equal to one in the black-disc (BD) limit.

We found [8] that interaction of  $q\bar{q}$  dipoles with transverse size  $\sim 0.3$  fm corresponding to  $Q^2 \sim 4\text{GeV}^2$  is still rather far from the BD regime for the range covered by HERA even for small impact parameters,  $b$ . At the same time a much stronger interaction in the gluon channel (a factor of 9/4 larger  $F^2$  in Eq. (3)) leads to  $\Gamma_{gg}(d \sim 0.3\text{fm}, x \sim 10^{-4})$  close to one in a large range of  $b$ , see Fig. 2. Proximity of  $\Gamma_{gg}$  to one in a wide range of  $b$  for  $Q^2 \sim 4\text{GeV}^2$  naturally explains a large probability of diffraction ( $\sim 30 \div 40\%$ ) in the gluon induced hard interactions which can be inferred from the HERA DGLAP analyses of the inclusive DIS diffractive data (see discussion and references in [3]).

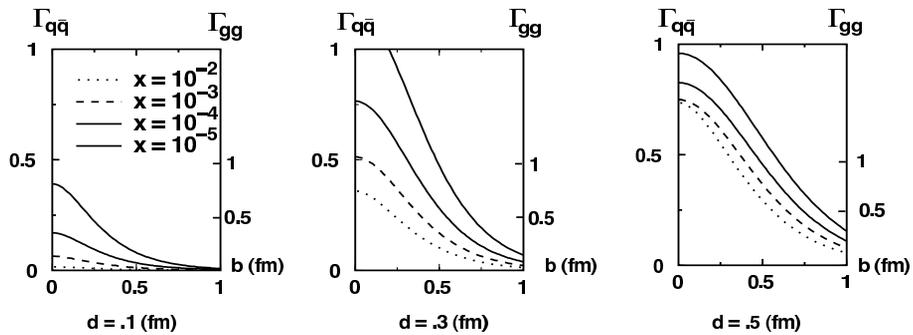


Fig. 2: The profile function of dipole-nucleon scattering,  $\Gamma^{dp}$ , as a function of the impact parameter,  $b$ , for various values of the dipole size,  $d$ , and  $x$ , as obtained from a phenomenological estimate outlined in the text. Shown are the results for  $q\bar{q}$  (left scale) and  $gg$  dipoles (right scale)

In the BD regime parton obtains transverse momenta of the order of the maximal  $p_t$  scale at which interaction remains black and also loses a substantial fraction of its longitudinal momentum (one can also think of this as a post selection of configurations in the incoming wave function with large transverse momenta; the simplest example is scattering of virtual photon in the BD regime [9]). The analysis of the data obtained by the BRAHMS [10] and STAR collaborations [11] on the leading pion production in the deuteron - gold collisions including forward - central rapidity correlations supports presence of this phenomenon for gluon densities comparable to those encountered at HERA [12].

At the LHC energies for the fragmentation region BD regime extends to quite large  $p_t$  for the leading partons (especially for gluons) up to  $\rho \sim 0.5\text{fm}$  which give important contribution to the central  $pp$  collisions (see Fig. 3 adapted from [13]).

Hence, in the  $pp$  collisions large  $x$  partons of nucleon "1" passing at small transverse distances  $\rho$  from the nucleon "2" should get large transverse momenta and also lose significant

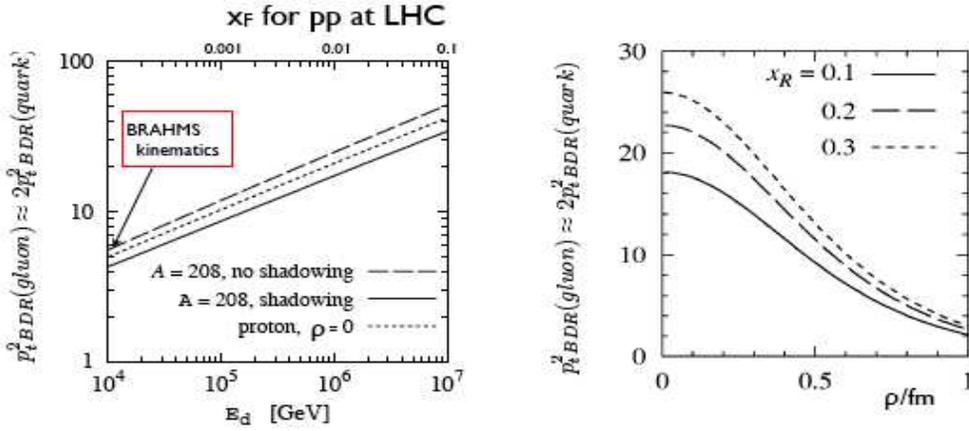


Fig. 3: Dependence of the maximum  $p_t^2$  for gluon for which interaction is close to the BD regime as a function of  $x_F$  (energy of the parton) for  $\rho = 0$  and as function of  $\rho$  for different  $x_R$  of the gluon for the LHC  $pp$  collisions.

fraction of energy. Note here that this effect is masked in many current MC event generators for  $pp$  collisions at the LHC, where a cutoff on minimal momentum transfer of the order 3 GeV is introduced.

One should note here that the necessity to tame intensity of hard collisions in  $pp$  scattering could be derived without invoking a study of the multiplicities of the produced hadrons as it is done e.g. in PYTHIA [4]. Instead, one can study the probability of inelastic interaction as a function of  $b$  which can be determined from unitarity - information on the elastic amplitude, and calculating the inelasticity due to hard parton-parton interactions. We found that for  $b \sim 1.5 fm$  (where uncertainties due to the contribution of multiparton interactions appear to be small) one needs to introduce a cutoff of the order of three GeV in order to avoid a contradiction with the S-channel unitarity [14]. The taming of the small  $x$  parton densities in the relevant  $x \geq 10^{-4}$  range for  $\rho \sim 0.7 fm$  is very small. Hence, it is not clear so far what dynamical mechanism is responsible for resolving problems with S-channel unitarity.

Modifications of the pattern of the collisions due to the large scale of BD regime for small  $\rho$  should be pronounced most prominently in the collisions at small impact parameters. Therefore they are enhanced in the processes of production of new particles which correspond to significantly smaller impact parameters than the minimum bias inelastic collisions. Among the expected effects are suppression of the leading baryon production, energy flow from forward region to smaller rapidities, larger central multiplicity, etc.

### 1.3 Centrality trigger for pp collisions

To study effects of high gluon densities it is desirable to develop a trigger for centrality in  $pp$  collisions [15]. We explore the observation that the leading nucleons are usually produced when number of "wounded" quarks,  $N_w$  is  $\leq 1$ . If  $N_w \geq 2$ , at least two quarks receive large transverse momenta they cannot combine into a leading nucleon as they fragment independently, so the spectra for  $N_w = 2$  and  $N_w = 3$  should be rather similar and shifted to much smaller  $x_F$  than in

soft interactions where the spectra of nucleons are known to be flat in  $x_F$  in a wide range of  $x_F$ .

We developed a MC event generator to quantify this observation. At the first step three quark configuration in one nucleon is generated with transverse coordinates given by the nucleon wave function. For given  $b$  we determine the gluon density encountered by each quark and if the gluon density corresponds to the BD regime, generate a transverse momentum for a quark using the model of [16] (we neglect the fractional energy losses expected in the BD regime [12]).

We implemented the fragmentation of the system produced in the first stage by constructing strings which decay using the LUND method. There are always two strings, drawn between a quark and a diquark from the interacting particles. When a quark of the diquark receives a high transverse momentum, the diquark becomes a system of two quarks and a junction. This has the nice property that one recovers the diquark when the invariant mass between the two quarks is small. The results are in good agreement with the qualitative expectation that spectra for  $N_w = 0$  and  $N_w = 1$  are similar and much harder than for  $N_w = 2, 3$  which are very similar, see Fig. 3 in Ref. [15].

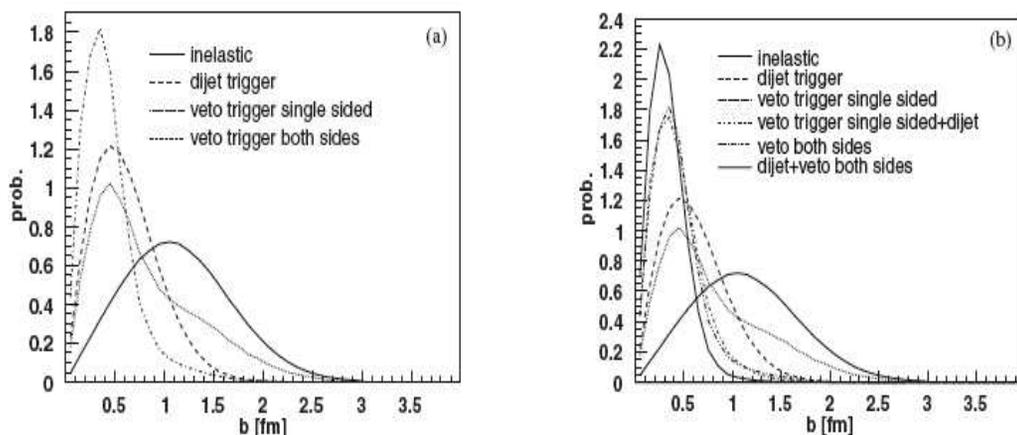


Fig. 4: (a):The combination of dijet and veto trigger gives the best constraints on central events in  $pp$ -collisions. (b):Impact parameter distributions for inelastic events, the dijet trigger and single and double sided veto-trigger (no baryon in the region  $x_F > 0.1$ ).

We find that  $N_w$  strongly depends on  $b$  with  $N_w \geq 2$  dominating for  $b \leq 0.5$ fm. A strong correlation of  $N_w$  with the multiplicity of leading baryons allows one to determine the effectiveness of a centrality trigger based on a veto for the production of leading baryons with  $x > x_{tr}$  as a function of  $x_{tr}$ . We find than an optimal value of  $x_{tr}$  is  $\sim 0.1$ . Current configurations of several LHC detectors allow to veto neutron production in this  $x$ -range. TOTEM, in addition, allows to veto production of protons with  $x_F > 0.8$ . Since neutron and proton multiplicities are similar, a one side veto for production of both charged and neutral baryons leads approximately to the same result as a two side veto for neutron production. Accordingly we will give results both for single side veto and for two side veto for both neutral and charged baryons (understanding that the full implementation of the latter option would require certain upgrades of the detectors some of which are currently under discussion). The results of the calculations are presented

in Fig. 4a together with the distribution over  $b$  for generic inelastic events and the central dijet trigger [13]. We see that the single side veto trigger leads to a centrality similar to that of a dijet trigger, while a double side veto leads to the most narrow distribution in  $b$ . An easy way to check this expectation would be to compare other characteristics of these types of events - one expects for example a progressive increase of the central multiplicity with a decrease of average  $b$ .

The most narrow distributions can be achieved by selecting events with dijets and without leading baryons, Fig. 4b in this case we reach the limit that  $\langle \rho_{tr} \rangle = (\langle \rho^2 \rangle + \langle b^2 \rangle)^{1/2}$  becomes comparable to  $\langle \rho \rangle$  which is the smallest possible average  $\langle \rho \rangle$  for  $pp$  or DIS collisions.

## 1.4 Conclusions

Understanding of the complexity of the nucleon structure is gradually emerging from the studies of hard interactions at HERA and Tevatron collider. In addition to revealing a small transverse localization of the gluon field one finds a number of other peculiarities: presence of significant fluctuations of the transverse size of the nucleon and the strength of the gluon fields, as well as indications of a lumpy structure of nucleon at low scale (constituent quarks).

Due to proximity of BD regime for a large range of virtualities the small  $x$  physics appears to be an unavoidable component of the new particle physics production at LHC.

One of the biggest challenges is to understand the mechanism and pattern of taming of parton interactions at transverse momenta of few GeV and how it affects spectra of leading partons in the central collisions. It maybe the best to study these phenomena using centrality triggers to amplify these phenomena. Among most sensitive tools are long range correlations in rapidity - central and forward hadron production, forward - backward correlations, transverse distribution in various hard processes with centrality trigger, etc. Large rapidity coverage of ATLAS and CMS / TOTEM allows to study correlations at much larger rapidity intervals than it was possible at previous colliders.

## 2 In-medium QCD and Cherenkov gluons vs. Mach waves at LHC

*Author: Igor M. Dremin*

The properties and evolution of the medium formed in ultrarelativistic heavy-ion collisions are widely debated. At the simplest level it is assumed to consist of a set of current quarks and gluons. The collective excitation modes of the medium may, however, play a crucial role. One of the ways to gain more knowledge about the excitation modes is to consider the propagation of relativistic partons through this matter. Phenomenologically their impact would be described by the nuclear permittivity of the matter corresponding to its response to passing partons. This approach is most successful for electro-dynamical processes in matter. Therefore, it is reasonable to modify the QCD equations by taking into account collective properties of the quark-gluon medium [17]. Strangely enough, this was not done earlier. For the sake of simplicity we consider here the gluodynamics only.

The classical lowest-order solution of these equations coincides with Abelian electro-dynamical results up to a trivial color factor. One of the most spectacular of them is Cherenkov ra-

diation and its properties. Now, Cherenkov gluons take the place of Cherenkov photons [18–20]. Their emission in high-energy hadronic collisions is described by the same formulae but with the nuclear permittivity in place of the usual one. Actually, one considers them as quasiparticles, i.e. quanta of the medium excitations leading to shock waves with properties determined by the permittivity.

Another problem of this approach is related to the notion of the rest system of the medium. It results in some specific features of this effect at LHC energies.

To begin, let us recall the classical in-vacuum Yang-Mills equations

$$D_\mu F^{\mu\nu} = J^\nu, \quad F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu - ig[A^\mu, A^\nu], \quad (6)$$

where  $A^\mu = iA_a^\mu T_a$ ;  $A_a(A_a^0 \equiv \Phi_a, \mathbf{A}_a)$  are the gauge field (scalar and vector) potentials, the color matrices  $T_a$  satisfy the relation  $[T_a, T_b] = if_{abc}T_c$ ,  $D_\mu = \partial_\mu - ig[A_\mu, \cdot]$ ,  $J^\nu(\rho, \mathbf{j})$  is a classical source current, and the metric is given by  $g^{\mu\nu} = \text{diag}(+, -, -, -)$ .

In the covariant gauge  $\partial_\mu A^\mu = 0$  they are written

$$\square A^\mu = J^\mu + ig[A_\nu, \partial^\nu A^\mu + F^{\mu\nu}], \quad (7)$$

where  $\square$  is the d'Alembertian operator.

The chromoelectric and chromomagnetic fields are  $E^\mu = F^{\mu 0}$ ,  $B^\mu = -\frac{1}{2}\epsilon^{\mu ij}F^{ij}$  or, as functions of the gauge potentials in vector notation,

$$\mathbf{E}_a = -\text{grad}\Phi_a - \frac{\partial \mathbf{A}_a}{\partial t} + gf_{abc}\mathbf{A}_b\Phi_c, \quad \mathbf{B}_a = \text{curl}\mathbf{A}_a - \frac{1}{2}gf_{abc}[\mathbf{A}_b\mathbf{A}_c]. \quad (8)$$

Herefrom, one easily rewrites the in-vacuum equations of motion (6) in vector form. We do not show them explicitly here (see [17]) and write down the equations of the in-medium gluon dynamics using the same method as in electrodynamics. We introduce the nuclear permittivity and denote it also by  $\epsilon$ , since this will not lead to any confusion. After that, one should replace  $\mathbf{E}_a$  by  $\epsilon\mathbf{E}_a$  and get

$$\epsilon(\text{div}\mathbf{E}_a - gf_{abc}\mathbf{A}_b\mathbf{E}_c) = \rho_a, \quad \text{curl}\mathbf{B}_a - \epsilon\frac{\partial \mathbf{E}_a}{\partial t} - gf_{abc}(\epsilon\Phi_b\mathbf{E}_c + [\mathbf{A}_b\mathbf{B}_c]) = \mathbf{j}_a. \quad (9)$$

The space-time dispersion of  $\epsilon$  is neglected here.

In terms of potentials these equations are cast in the form

$$\begin{aligned} \Delta \mathbf{A}_a - \epsilon \frac{\partial^2 \mathbf{A}_a}{\partial t^2} &= -\mathbf{j}_a - gf_{abc} \left( \frac{1}{2} \text{curl}[\mathbf{A}_b, \mathbf{A}_c] + \frac{\partial}{\partial t}(\mathbf{A}_b\Phi_c) + [\mathbf{A}_b \text{curl}\mathbf{A}_c] - \right. \\ &\quad \left. \epsilon\Phi_b \frac{\partial \mathbf{A}_c}{\partial t} - \epsilon\Phi_b \text{grad}\Phi_c - \frac{1}{2}gf_{cmn}[\mathbf{A}_b[\mathbf{A}_m\mathbf{A}_n]] + g\epsilon f_{cmn}\Phi_b\mathbf{A}_m\Phi_n \right), \end{aligned} \quad (10)$$

$$\begin{aligned} \Delta \Phi_a - \epsilon \frac{\partial^2 \Phi_a}{\partial t^2} &= -\frac{\rho_a}{\epsilon} + gf_{abc}(2\mathbf{A}_c \text{grad}\Phi_b + \mathbf{A}_b \frac{\partial \mathbf{A}_c}{\partial t} + \frac{\partial \Phi_b}{\partial t} \Phi_c) - \\ &\quad g^2 f_{amn} f_{nlb} \mathbf{A}_m \mathbf{A}_l \Phi_b. \end{aligned} \quad (11)$$

If the terms with coupling constant  $g$  are omitted, one gets the set of Abelian equations, that differ from electrodynamical equations by the color index  $a$  only. The external current is due to a parton moving fast relative to partons "at rest".

The crucial distinction between (7) and (10), (11) is that there is no radiation (the field strength is zero in the forward light-cone and no gluons are produced) in the lowest order solution of (7), and it is admitted for (10), (11), because  $\epsilon$  takes into account the collective response (color polarization) of the nuclear matter.

Cherenkov effects are especially suited for treating them by classical approach to (10), (11). Their unique feature is independence of the coherence of subsequent emissions on the time interval between these processes. The lack of balance of the phase  $\Delta\phi$  between emissions with frequency  $\omega = k/\sqrt{\epsilon}$  separated by the time interval  $\Delta t$  (or the length  $\Delta z = v\Delta t$ ) is given by

$$\Delta\phi = \omega\Delta t - k\Delta z \cos\theta = k\Delta z\left(\frac{1}{v\sqrt{\epsilon}} - \cos\theta\right) \quad (12)$$

up to terms that vanish for large distances. For Cherenkov effects the angle  $\theta$  is

$$\cos\theta = \frac{1}{v\sqrt{\epsilon}}. \quad (13)$$

The coherence condition  $\Delta\phi = 0$  is valid independent of  $\Delta z$ . This is a crucial property specific for Cherenkov radiation only. The fields  $(\Phi_a, \mathbf{A}_a)$  and the classical current for in-medium gluon dynamics can be represented by the product of the electrodynamical expressions  $(\Phi, \mathbf{A})$  and the color matrix  $T_a$ .

Let us recall the Abelian solution for the current with velocity  $\mathbf{v}$  along  $z$ -axis:

$$\mathbf{j}(\mathbf{r}, t) = \mathbf{v}\rho(\mathbf{r}, t) = 4\pi g\mathbf{v}\delta(\mathbf{r} - \mathbf{v}t). \quad (14)$$

In the lowest order the solutions for the scalar and vector potentials are related  $\mathbf{A}^{(1)}(\mathbf{r}, t) = \epsilon\mathbf{v}\Phi^{(1)}(\mathbf{r}, t)$  and

$$\Phi^{(1)}(\mathbf{r}, t) = \frac{2g}{\epsilon} \frac{\theta(vt - z - r_{\perp}\sqrt{\epsilon v^2 - 1})}{\sqrt{(vt - z)^2 - r_{\perp}^2(\epsilon v^2 - 1)}}. \quad (15)$$

Here  $r_{\perp} = \sqrt{x^2 + y^2}$  is the cylindrical coordinate;  $z$  symmetry axis. The cone

$$z = vt - r_{\perp}\sqrt{\epsilon v^2 - 1} \quad (16)$$

determines the position of the shock wave due to the  $\theta$ -function in (15). The field is localized within this cone and decreases with time as  $1/t$  at any fixed point. The gluons emission is perpendicular to the cone (16) at the Cherenkov angle (13).

Due to the antisymmetry of  $f_{abc}$ , the higher order terms ( $g^3, \dots$ ) are equal to zero for any solution multiplicative in space-time and color as seen from (10), (11).

The expression for the intensity of the radiation is given by the Tamm-Frank formula (up to Casimir operators) that leads to infinity for constant  $\epsilon$ . The  $\omega$ -dependence of  $\epsilon$  (dispersion), its imaginary part (absorption) and chromomagnetic permeability can be taken into account [17].

The attempts to calculate the nuclear permittivity from first principles are not very convincing. It can be obtained from the polarization operator. The corresponding dispersion branches have been computed in the lowest order perturbation theory [21, 22]. The properties of collective excitations have been studied in the framework of the thermal field theories (see, e.g., [23]). The results with an additional phenomenological ad hoc assumption about the role of resonances were used in a simplified model of scalar fields [20] to show that the nuclear permittivity can be larger than 1, i.e. admits Cherenkov gluons. Extensive studies were performed in [24]. No final decision about the nuclear permittivity is yet obtained from these approaches. It must be nontrivial problem because we know that, e.g., the energy dependence of the refractive index of water [25] (especially, its imaginary part) is so complicated that it is not described quantitatively in electrodynamics.

Therefore, we prefer to use the general formulae of the scattering theory to estimate the nuclear permittivity. It is related to the refractive index  $n$  of the medium  $\epsilon = n^2$  and the latter one is expressed through the real part of the forward scattering amplitude of the refracted quanta  $\text{Re}F(0^\circ, E)$  by

$$\text{Re}n(E) = 1 + \Delta n_R = 1 + \frac{6m_\pi^3\nu}{E^2}\text{Re}F(E) = 1 + \frac{3m_\pi^3\nu}{4\pi E}\sigma(E)\rho(E). \quad (17)$$

Here  $E$  denotes the energy,  $\nu$  the number of scatterers within a single nucleon,  $m_\pi$  the pion mass,  $\sigma(E)$  the cross section and  $\rho(E)$  the ratio of real to imaginary parts of the forward scattering amplitude  $F(E)$ .

Thus the emission of Cherenkov gluons is possible only for processes with positive  $\text{Re}F(E)$  or  $\rho(E)$ . Unfortunately, we are unable to calculate directly in QCD these characteristics of gluons and have to rely on analogies and our knowledge of the properties of hadrons. The only experimental facts we get for this medium are brought about by particles registered at the final stage. They have some features in common, which (one may hope!) are also relevant for gluons as the carriers of the strong forces. Those are the resonant behavior of amplitudes at rather low energies and the positive real part of the forward scattering amplitudes at very high energies for hadron-hadron and photon-hadron processes as measured from the interference of the Coulomb and hadronic parts of the amplitudes.  $\text{Re}F(0^\circ, E)$  is always positive (i.e.,  $n > 1$ ) within the low-mass wings of the Breit-Wigner resonances. This shows that the necessary condition for Cherenkov effects  $n > 1$  is satisfied at least within these two energy intervals. This fact was used to describe experimental observations at SPS, RHIC and cosmic ray energies. The asymmetry of the  $\rho$ -meson shape at SPS [26] and azimuthal correlations of in-medium jets at RHIC [27–30] were explained by emission of comparatively low-energy Cherenkov gluons [31, 32]. The parton density and intensity of the radiation were estimated. In its turn, cosmic ray data [33] at energies corresponding to LHC require very high-energy gluons to be emitted by the ultrarelativistic partons moving along the collision axis [18, 19]. Let us note the important difference from electrodynamics, where  $n < 1$  at high frequencies.

The in-medium equations are not Lorentz-invariant. There is no problem in macroscopic electrodynamics, because the rest system of the macroscopic matter is well defined and its permittivity is considered there. For collisions of two nuclei (or hadrons) it requires special discussion.

Let us consider a particular parton that radiates in the nuclear matter. It would "feel"

the surrounding medium at rest if the momenta of all other partons, with which this parton can interact, are smaller and sum to zero. In RHIC experiments the triggers, that registered the jets (created by partons), were positioned at  $90^\circ$  to the collision axis. Such partons should be produced by two initial forward-backward moving partons scattered at  $90^\circ$ . The total momentum of the other partons (medium spectators) is balanced, because for such a geometry the partons from both nuclei play the role of spectators forming the medium. Thus the center of mass system is the proper one to consider the nuclear matter at rest in this experiment. The permittivity must be defined there. The Cherenkov rings consisting of hadrons have been registered around the away-side jet, which traversed the nuclear medium. This geometry requires, however, high statistics, because the rare process of scattering at  $90^\circ$  has been chosen.

The forward (backward) moving partons are much more numerous and have higher energies. However, one cannot treat the radiation of such a primary parton in the c.m.s. in a similar way, because the momentum of the spectators is different from zero, i.e. the matter is not at rest. Now the spectators (the medium) are formed from the partons of another nucleus only. Then the rest system of the medium coincides with the rest system of that nucleus and the permittivity should refer to this system. The Cherenkov radiation of such highly energetic partons must be considered there. That is what was done for interpretation of the cosmic ray event in [18, 19]. This discussion shows that one must carefully define the rest system for other geometries of the experiment with triggers positioned at different angles.

Thus our conclusion is that the definition of  $\epsilon$  depends on the geometry of the experiment. Its corollary is that partons moving in different directions with different energies can "feel" different states of matter in the **same** collision of two nuclei because of the permittivity dispersion. The transversely scattered partons with comparatively low energies can analyze the matter with rather large permittivity corresponding to the resonance region, while the forward moving partons with high energies would "observe" a low permittivity in the same collision. This peculiar feature can help scan the  $(\ln x, Q^2)$ -plane as discussed in [34]. It explains also the different values of  $\epsilon$  needed for the description of the RHIC and cosmic ray data.

These conclusions can be checked at LHC, because both RHIC and cosmic ray geometry will become available there. The energy of the forward moving partons would exceed the thresholds above which  $n > 1$ . Then both types of experiments can be done, i.e. the  $90^\circ$ -trigger and non-trigger forward-backward partons experiments. The predicted results for  $90^\circ$ -trigger geometry are similar to those at RHIC. The non-trigger Cherenkov gluons should be emitted within the rings at polar angles of tens degrees in c.m.s. at LHC by the forward moving partons (and symmetrically by the backward ones) according to some events observed in cosmic rays [32, 33].

Let us compare the conclusions for Cherenkov and Mach shock waves. The Cherenkov gluons are described as the transverse waves while the Mach waves are longitudinal. Up to now, no experimental signatures of these features were proposed.

The most important experimental fact is the position of the maxima of humps in two-particle correlations. They are displaced from the away-side jet by 1.05-1.23 radian [35–38]. This requires rather large values of  $\text{Re}\epsilon \sim 2 - 3$  and indicates high density of the medium [32] that agrees with other conclusions. The fits of the humps with complex permittivity are in progress. The maxima due to Mach shock waves should be shifted by the smaller value 0.955 if the relativistic equation of state is used ( $\cos \theta = 1/\sqrt{3}$ ). To fit experimental values one must

consider different equation of state. In three-particle correlations, this displacement is about 1.38 [27–29].

There are some claims [27–30] that Cherenkov effect contradicts to experimental observations because it predicts the shift of these maxima to smaller angles for larger momenta. They refer to the prediction made in [20]. However, the conclusions of this paper about the momentum dependence of the refractive index can hardly be considered as quantitative ones because the oversimplified scalar  $\Phi^3$ -model with simplest resonance insertions was used for computing the refractive index. In view of difficult task of its calculation discussed above, the fits of maxima seem to be more important for our conclusions about the validity of the two schemes.

Mach waves should appear for forward moving partons at RHIC but were not found. The energy threshold of  $\epsilon$  explains this phenomenon for Cherenkov gluons.

## References

- [1] S. J. Brodsky, L. Frankfurt, J. F. Gunion, A. H. Mueller, and M. Strikman, *Phys. Rev.* **D50**, 3134 (1994), [arXiv:hep-ph/9402283](#).
- [2] J. C. Collins, L. Frankfurt, and M. Strikman, *Phys. Rev.* **D56**, 2982 (1997), [arXiv:hep-ph/9611433](#).
- [3] L. Frankfurt, M. Strikman, and C. Weiss, *Ann. Rev. Nucl. Part. Sci.* **55**, 403 (2005), [arXiv:hep-ph/0507286](#).
- [4] T. Sjostrand, S. Mrenna, and P. Skands, *JHEP* **05**, 026 (2006), [arXiv:hep-ph/0603175](#).
- [5] CDF Collaboration, F. Abe *et al.*, *Phys. Rev. Lett.* **79**, 584 (1997).
- [6] CDF Collaboration, F. Abe *et al.*, *Phys. Rev.* **D56**, 3811 (1997).
- [7] L. Frankfurt, M. Strikman, D. Treleani, and C. Weiss, *Phys. Rev. Lett.* **101**, 202003 (2008), [arXiv:0808.0182 \[hep-ph\]](#).
- [8] T. Rogers, V. Guzey, M. Strikman, and X. Zu, *Phys. Rev.* **D69**, 074011 (2004), [arXiv:hep-ph/0309099](#).
- [9] L. Frankfurt, V. Guzey, M. McDermott, and M. Strikman, *Phys. Rev. Lett.* **87**, 192301 (2001), [arXiv:hep-ph/0104154](#).
- [10] BRAHMS Collaboration, I. Arsene *et al.*, *Phys. Rev. Lett.* **93**, 242303 (2004), [arXiv:nucl-ex/0403005](#).
- [11] STAR Collaboration, J. Adams *et al.*, *Phys. Rev. Lett.* **97**, 152302 (2006), [arXiv:nucl-ex/0602011](#).
- [12] L. Frankfurt and M. Strikman, *Phys. Lett.* **B645**, 412 (2007).
- [13] L. Frankfurt, M. Strikman, and C. Weiss, *Phys. Rev.* **D69**, 114010 (2004), [arXiv:hep-ph/0311231](#).

- [14] T. C. Rogers, A. M. Stasto, and M. I. Strikman, Phys. Rev. **D77**, 114009 (2008), arXiv:0801.0303 [hep-ph].
- [15] H. J. Drescher and M. Strikman, Phys. Rev. Lett. **100**, 152002 (2008).
- [16] D. Boer and A. Dumitru, Phys. Lett. **B556**, 33 (2003), arXiv:hep-ph/0212260.
- [17] I. M. Dremin, Eur. Phys. J. **C56**, 81 (2008), arXiv:0802.4022 [hep-ph].
- [18] I. M. Dremin, JETP Lett. **30**, 140 (1979).
- [19] I. M. Dremin, Sov. J. Nucl. Phys. **33**, 726 (1981).
- [20] V. Koch, A. Majumder, and X.-N. Wang, Phys. Rev. Lett. **96**, 172302 (2006), arXiv:nucl-th/0507063.
- [21] O. K. Kalashnikov and V. V. Klimov, Sov. J. Nucl. Phys. **31**, 699 (1980).
- [22] H. A. Weldon, Phys. Rev. **D26**, 1394 (1982).
- [23] J.-P. Blaizot and E. Iancu, Phys. Rept. **359**, 355 (2002), arXiv:hep-ph/0101103.
- [24] M. K. Djongolov, S. Pisov, and V. Rizov, J. Phys. **G30**, 425 (2004), arXiv:hep-ph/0303141.
- [25] J. D. Jackson, *Classical electrodynamics*. John Wiley and Sons, 1998.
- [26] NA60 Collaboration, R. Arnaldi *et al.*, Phys. Rev. Lett. **96**, 162302 (2006), arXiv:nucl-ex/0605007.
- [27] J. G. Ulery, PoS **LHC07**, 036 (2007), arXiv:0709.1633 [nucl-ex].
- [28] J. G. Ulery (2008), arXiv:0801.4904 [nucl-ex].
- [29] J. G. Ulery (2008), arXiv:0807.1613 [nucl-ex].
- [30] PHENIX Collaboration, N. N. Ajitanand, Nucl. Phys. **A783**, 519 (2007), arXiv:nucl-ex/0609038.
- [31] I. M. Dremin and V. A. Nechitailo (2007), arXiv:0704.1081 [hep-ph].
- [32] I. M. Dremin, Nucl. Phys. **A767**, 233 (2006), arXiv:hep-ph/0507167.
- [33] A. V. Apanasenko, N. A. Dobrotin, I. M. Dremin, and K. A. Kotelnikov, JETP Lett. **30**, 145 (1979).
- [34] I. M. Dremin, Int. J. Mod. Phys. **A22**, 3087 (2007), arXiv:0706.0596 [hep-ph].
- [35] STAR Collaboration, K. H. Ackermann *et al.*, Phys. Rev. Lett. **86**, 402 (2001), arXiv:nucl-ex/0009011.

- [36] STAR Collaboration, F. Wang, J. Phys. **G30**, S1299 (2004),  
arXiv:nucl-ex/0404010.
- [37] PHENIX Collaboration, A. Adare *et al.*, Phys. Rev. **C77**, 011901 (2008),  
arXiv:0705.3238 [nucl-ex].
- [38] PHENIX Collaboration, A. Adare *et al.*, Phys. Rev. **C78**, 014901 (2008),  
arXiv:0801.4545 [nucl-ex].

## Chapter 3

# Working Group Heavy Quarks (Charm and Beauty)

### Convenors:

*M. Cacciari (Paris VI & VII.),*

*A. Dainese (INFN, ALICE),*

*A. Geiser (DESY, ZEUS)*

*H. Spiesberger (U. Mainz),*

*Contactpersons: K. Lipka (U. Hamburg, H1), Ulrich Uwer (CERN)*

# Introduction

*M. Cacciari<sup>a</sup>, A. Dainese<sup>b</sup>, A. Geiser<sup>c</sup>, H. Spiesberger<sup>d</sup>*

<sup>a</sup> LPTHE, UPMC Paris

<sup>b</sup> INFN Legnaro

<sup>c</sup> Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany

<sup>d</sup> University Mainz, Germany

This document is a collection of contributions to the series of workshops, having taken place during the years 2006 – 2008 at CERN and DESY, on aspects of heavy quark physics relevant at the transition from the HERA to the LHC eras of experimentation. In three sections we review recent experimental results from HERA, describe the plans for coming analyses at the LHC and collect various reports about new results from theoretical work. The present report extends the proceedings of a previous workshop which are available online <sup>1</sup> and contain also a general theoretical review of various approaches in heavy quark production to which we refer for further reference.

The first section below reviews recent measurements of charm and beauty production in *ep* collisions at HERA. Heavy quark tagging methods used by the ZEUS and H1 experiments are described. Cross section results in both photoproduction and deep inelastic scattering are compared with NLO QCD predictions. In general the data are well described by the calculations. Studies of charm fragmentation yield compatibility with the assumption of universality at large transverse momenta, but illustrate some problems with this assumption in the threshold region. The DIS cross sections receive large contributions from the charm and beauty content of the proton,  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$ . The corresponding most recent measurements are compared to next-to-leading order QCD predictions using different parameterisations of the theory, and of the gluon density in the proton.

The tests of the fragmentation function, the gluon density, and details of the theoretical treatment of the charm and beauty masses performed on the basis of HERA experimental data are of direct interest to corresponding applications at the LHC. In section two, after reviewing the main heavy flavour results from experiments at the Relativistic Heavy Ion Collider (RHIC), we present the expected performance for some of the most significant measurements in the heavy flavour sector at the Large Hadron Collider (LHC), for the experiments ALICE, ATLAS, and CMS.

A proper inclusion of heavy quark mass effects in parton distribution function fits has proved crucial. In the theory part of this document, we present a review of these effects in DIS and their impact on global analyses and lay out all elements of a properly defined general mass variable flavor number scheme (GM VFNS) that are shared by all modern formulations of the problem. We also report about progress in a number of theoretical problems related to exclusive measurements of heavy flavors. These topics include fragmentation functions for charmed mesons including finite mass effects, fragmentation functions including non-perturbative corrections based on an effective QCD coupling, a discussion of the status of higher-order calculations

---

<sup>1</sup><http://www.desy.de/~heralhc/proceedings/proceedings.html>

for top quark production and for polarized structure functions, heavy quark and quarkonium production in the Regge limit, double heavy baryon production, tests of time reversal and CP symmetry in  $\Lambda_b$  decays, as well as a study of the general properties of massive exotic hadrons that will be relevant for an understanding of their detection at the LHC.

# Experimental study of heavy flavour production at HERA

*S. Boutle<sup>a</sup>, J. Bracinik<sup>b</sup>, A. Geiser<sup>c</sup>, G. Grindhammer<sup>d</sup>, A.W. Jung<sup>e</sup>, P. Roloff<sup>c,f</sup>, Z. Ruriková<sup>c</sup>, M. Turcato<sup>f</sup>, A. Yagües-Molina<sup>c</sup>*

<sup>a</sup> University College London, Gower Street, London WC1E 6BT, United Kingdom

<sup>b</sup> School of Physics and Astronomy, University of Birmingham - UK

<sup>c</sup> Deutsches Elektronen-Synchrotron DESY, Notkestraße 85, 22607 Hamburg, Germany

<sup>d</sup> Max-Planck-Institut f. Physik, Werner-Heisenberg-Institut Muenchen, Germany

<sup>e</sup> University of Heidelberg - Kirchhoff-Institute for Physics, Im Neuenheimer Feld 227, 69120 Heidelberg, Germany

<sup>f</sup> Universität Hamburg, Institut für Experimentalphysik, Luruper Chaussee 149, 22761 Hamburg, Germany

## Abstract

Recent measurements of charm and beauty production in  $ep$  collisions at HERA are reviewed. Heavy quark tagging methods used by the ZEUS and H1 experiments are described. Cross section results in both photoproduction and deep inelastic scattering (DIS) are compared with NLO QCD predictions. In general the data are well described by the calculations. Studies of charm fragmentation yield compatibility with the assumption of universality at large transverse momenta, but illustrate some problems with this assumption in the threshold region. The DIS cross sections can also be expressed in terms of the charm and beauty content of the proton,  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$ . The most recent measurements are compared to next-to-leading order QCD predictions using different parameterisations of the theory, and of the gluon density in the proton. The tests of the fragmentation function, the gluon density, and details of the theoretical treatment of the charm and beauty masses are of direct interest to corresponding applications at the LHC.

*Coordinator/editor: A. Geiser*

## 1 Charm production at HERA: Experimental overview

*Author: A. Jung*

### 1.1 Introduction

Several new measurements of open charm production have been performed by the H1 collaboration

- $D^{*\pm}$  Production at low  $Q^2$  with the H1 Detector [1, 2]
- Measurement of the  $D^{*\pm}$  Production cross section in Photoproduction with the H1 Detector using HERA II data [3]
- Study of Charm Fragmentation into  $D^{*\pm}$  Mesons in Deep-Inelastic Scattering at HERA [4]

and by the ZEUS collaboration

- $D^{*\pm}$  in DIS and Measurement of  $F_2^c$  [5]
- Measurement of  $D^{*\pm}$  Meson Production in DIS ep Scattering at low  $Q^2$  [6]
- Measurement of excited charm and charm-strange mesons production at HERA [7]
- Measurement of the charm fragmentation fractions [8] and fragmentation function [9]

The details of the measurements like the visible range will not be discussed here as they are given in the literature cited for each measurement. For both experiments high statistic charm event samples are tagged by  $D^{*\pm}$  mesons reconstructed in the golden decay channel:  $D^{*\pm} \rightarrow D^0 \pi_{\text{slow}}^\pm \rightarrow K^\mp \pi^\pm \pi_{\text{slow}}^\pm$ . The well known mass difference method reduces symmetric systematic uncertainties and allows the extraction of the  $D^*$  meson signal by fits out of the background dominated data samples. Other D mesons can also be identified via their characteristic mass peaks. In addition to that method charmed mesons like  $D^+$ ,  $D_s^+$  mesons are tagged via lifetime measurements from the high resolution silicon vertex detectors used by both experiments in HERAII.

The results on the fragmentation function will be discussed in section 2.

## 1.2 Results of open charm production

The H1 photoproduction analysis [3] makes use of the H1 Fast Track Trigger (FTT) [10, 11] which enhanced the capabilities of heavy flavor measurements at H1 by a selective on-line track based event reconstruction [12]. Due to these improvements the phase space and the available statistics of the measurement has been significantly enlarged compared to the previous H1 photoproduction analysis [13].

The large statistics allows precise double differential measurements. For the H1 photoproduction measurement the data are reasonably well described except for special regions of the phase space and correlations. Especially the  $W_{\gamma P}$  dependence is not described by the NLO prediction using the FFNS. The correlation between  $\eta$  and  $p_T$  as shown in Figure 1 (right) is compared to the NLO QCD predictions in the FFNS [14] and GM-VFNS; it turns out that the NLO QCD predictions are able to describe the correlation between  $\eta$  and  $p_T$  in photoproduction. Nevertheless the NLO predictions show an increasing deficit at forward  $\eta > 0$  which is largest at high  $p_T$  where the  $D^*$  data prefer the upper edge of the error band. For the photoproduction regime the relatively large theoretical uncertainty especially at small  $p_T$  arises from the scale variation. For comparison also the double differential measurement in DIS from H1 is shown on the left side of Figure 1. The H1 DIS analysis [1] uses the full HERAII luminosity. Because of the large statistics the analysis is almost everywhere dominated by the systematic error. The use of electron and hadron quantities combined in the  $e\Sigma$  reconstruction method [15] for the reconstruction of the kinematic variables allows lower inelasticities and smaller systematic uncertainties compared to previous H1 DIS analyses [2].

The measured single and double differential  $D^*$  production cross sections are in general well described by the next-to-leading order QCD predictions in the FFNS. The theoretical uncertainty of the predictions is dominated by the mass variation of the charm quark but is in general smaller than in photoproduction because of the additional scale  $Q^2$ . The small excess in data at forward directions  $\eta > 0$  (seen previously by H1 [2]) turns out to be located at low  $p_T$  as it can be seen

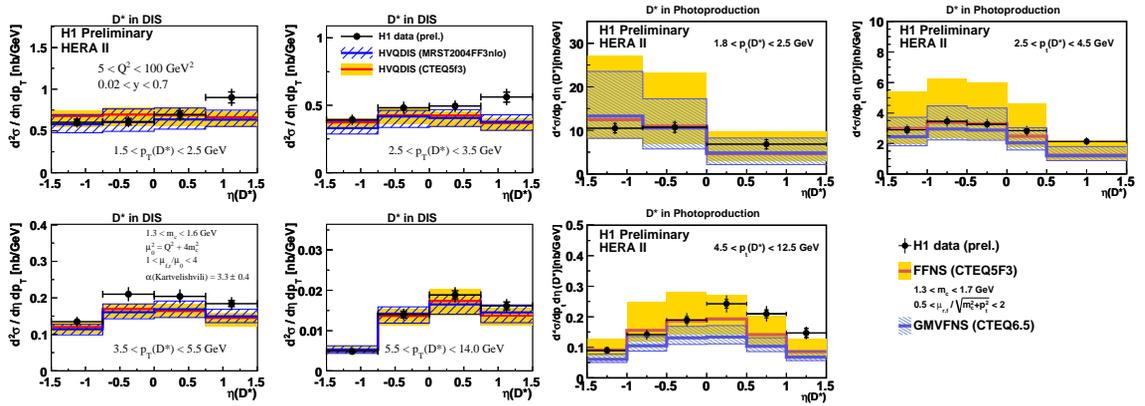


Fig. 1: The double differential cross section in  $\eta(D^*)$  and  $p_T(D^*)$  for the DIS (left) and photoproduction (right) regime compared to the NLO QCD predictions.

in comparison to the NLO prediction for the double differential distribution (see Figure 1 left). The data are above the predictions for the low  $p_T$  region at forward directions which is different to the photoproduction region where the data prefer the upper edge of the prediction at large  $p_T$ . The small discrepancy at forward directions can already be seen in the single differential  $\eta(D^*)$  distribution in comparison to the NLO QCD prediction as shown in Figure 2 left. The recent

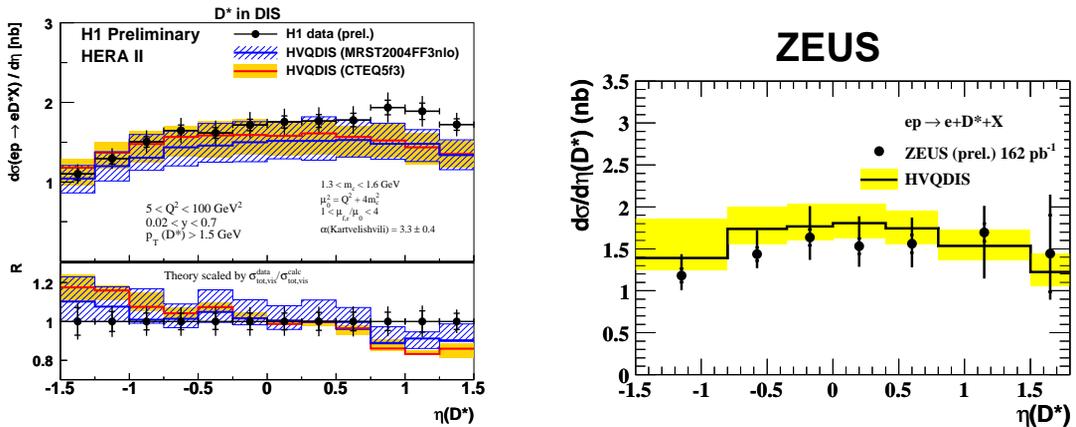


Fig. 2: The  $D^*$  cross section as a function of  $\eta(D^*)$  as measured by H1 using the luminosity of the whole HERAII data taking (left) and from ZEUS as measured using the luminosity of the years 2003 – 2005 (right).

$D^*$  measurement from the ZEUS collaboration [5] measures the same  $p_T(D^*)$  and  $y$  region but covers a slightly larger range in  $\eta(D^*)$  and goes up to larger  $Q^2$ . The ZEUS result (figure 2 right) is in good agreement with the NLO QCD predictions and in agreement with the result from H1 within errors.

A cross section measurement at very low  $Q^2$  for  $D^*$  production in DIS has been performed by ZEUS [6] using the beam pipe calorimeter. The overall  $Q^2$  range including this new measurement is shown in figure 3 (left) with a nice agreement to the NLO QCD prediction.

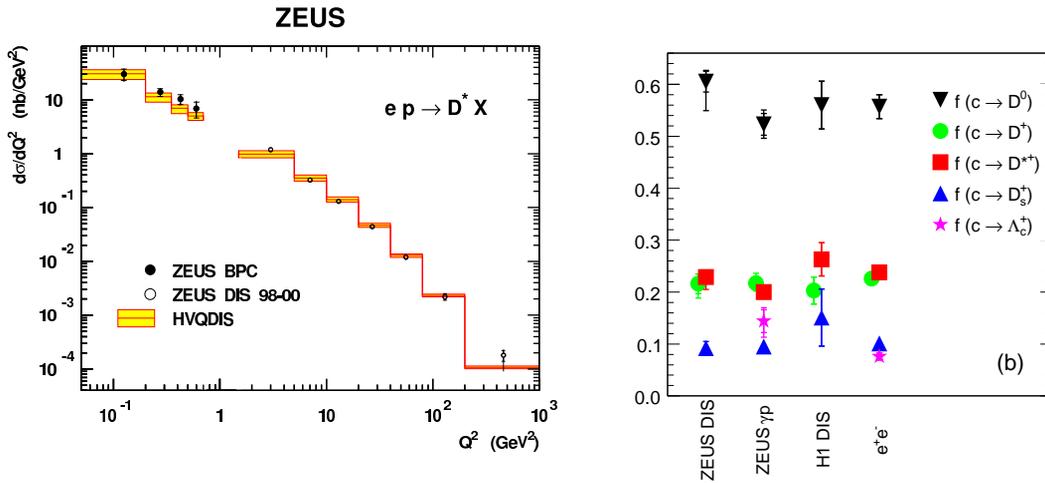


Fig. 3: The  $Q^2$  distribution including the new ZEUS measurement at very low  $Q^2$  shown right and the fragmentation fractions as measured by ZEUS and other experiment for various D mesons.

ZEUS has measured the charm production cross section of  $D^*$ ,  $D^0$ ,  $D^+$ ,  $D_s^+$  [8]. in order to determine the fragmentation fractions of charm into each meson. The charm fragmentation fractions as shown in figure 3 agree with the ones extracted by H1. Because of the agreement with the fragmentation fractions from  $e^+e^-$  the conclusion is that they do not depend on the hard subprocess and are in that sense universal.

In addition to the test of the QCD predictions in differential distributions another stringent test of QCD is possible since the  $D^*$  measurement involves the gluon density which drives the  $D^*$  production via the BGF production mechanism. Several approaches exist to measure the gluon density. The well established approach to measure the charm structure function will be covered elsewhere. In order to have an impact on the fits of the gluon density it is necessary that the cross section data have the highest possible precision. At present stage H1 and ZEUS enter the precision era of charm measurements where a single differential distribution has at least some sensitivity to the proton PDF, e.g. the  $\eta(D^*)$  distribution measured by H1 shows a better compatibility to the predictions if a proton PDF is used where a gluon density providing a less steep rise towards small  $x$  is used. However, the significance of the sensitivity is diminished by the relatively large theoretical uncertainties. The available  $D^*$  cross section data can also be used to fit the gluon density directly from the differential distributions in  $\eta(D^*)$ ,  $p_T(D^*)$ ,  $z(D^*)$ ,  $x$ ,  $Q^2$  [16].

In order to further increase the data precision it is possible to combine data from H1 and ZEUS on the basis of  $(D^*)$  cross sections or at the level of  $F_2^c$  extractions. At the level of  $F_2^c$  also the combination of data within one experiment from different  $F_2^c$  measurement methods, i.e. from  $D^*$  cross sections and from lifetime measurements, provides additional information.

### 1.3 Conclusions

At the present stage H1 and ZEUS enter the precision era of charm measurements with the large statistic of about  $0.5 \text{ fb}^{-1}$  per experiment provided by the HERAII running period. These data are currently analyzed and first preliminary results with high precision are available. In general the description by the next-to-leading order QCD predictions is reasonable except for some regions of the phase space. In order to get more insights and to have a significant impact on the fits of the PDFs the cross section data must be very precise and in addition cover the largest possible phase space. New results with such improvements are still expected to come.

## 2 Study of Charm Quark Fragmentation at HERA

*Authors: J. Bracinik, G. Grindhammer, Z. Rurikova*

The inclusive cross section for the production of heavy hadrons in  $ep$  collisions can be expressed as a convolution of three terms, describing the structure of the proton, the hard subprocess and the transition of partons to colourless heavy and light hadrons. The term describing the transition of partons to hadrons, also referred to as fragmentation function (FF), contains a non-perturbative component and thus must be experimentally determined. Charm quark fragmentation has been already extensively studied in  $e^+e^-$  annihilation experiments, and the parameters of various phenomenological models (i.e. the Lund string model, independent fragmentation model), which have been developed to describe the fragmentation process, have been tuned. By studying the charm quark fragmentation function also in  $ep$  collisions one can experimentally test, if the assumed universality of fragmentation functions, i.e. their portability from the calculation of processes in  $e^+e^-$  to processes in  $ep$  or  $pp/\bar{p}$ , really holds.

Since fragmentation functions describe the longitudinal momentum fraction transferred from the parton to the hadron they cannot be measured directly. The differential cross section as a function of suitably defined observables sensitive to the FF has to be measured and used to extract the parameters of the FF by comparing the data with the prediction of a given model.

The fragmentation of charm quarks into  $D^{*\pm}$  mesons in  $ep$  collisions has been studied by the H1 [17] and ZEUS [18] collaborations in both deep-inelastic scattering (DIS) and photoproduction (PHP), respectively. A so called hemisphere and a jet observable have been used by H1 and a jet observable by ZEUS.

In case of the jet observable, the momentum of the charm quark is approximated by the momentum of the reconstructed jet, which includes a  $D^{*\pm}$  meson, leading to the definition of  $z_{\text{jet}} = (E + P_L)_{D^*} / (E + P)_{\text{jet}}$ , where the longitudinal momentum  $P_{LD^*}$  is defined with respect to the three-momentum of the jet.

In case of the hemisphere observable, the kinematics of charm production, known to proceed mainly via photon-gluon fusion, is taken into account. In the  $\gamma^*p$  rest-frame the charm and anti-charm quarks are moving in the direction of the virtual photon (see figure 4 left), hence, the contributions from initial state radiation and the proton remnant can be strongly suppressed by discarding all particles with momenta pointing to the proton direction. Furthermore, since the transverse momenta of the charm quarks are balanced in this frame, the remaining particles may be divided into two hemispheres, one containing the fragmentation products of the charm quark, and the other one those of the anti-charm quark (see figure 4 right). This division into hemi-

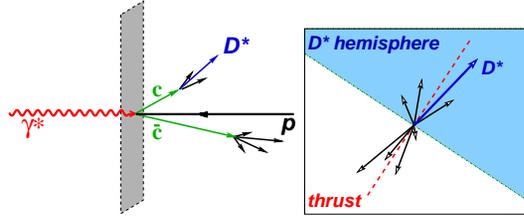


Fig. 4: Kinematics of charm/anticharm production in the  $\gamma^* p$  rest-frame as used for the definition of  $z_{\text{hem}}$ .

spheres is done by reconstructing the thrust axis in a plane perpendicular to the  $\gamma^* p$ -axis. The particles belonging to the same hemisphere as the  $D^{*\pm}$  meson are considered to be the products of the same quark and the sum of their four-momenta is used to approximate the four-momentum of the original quark, leading to this definition of  $z_{\text{hem}} = (E + P_L)_{D^*} / (E + P)_{\text{hem}}$ .

The ZEUS collaboration performed a measurement of the normalized differential cross section of  $D^{*\pm}$  meson production as a function of  $z_{\text{jet}}$  in photoproduction (kinematic range  $Q^2 < 1 \text{ GeV}^2$  and  $130 < W < 280 \text{ GeV}$ ). The  $D^{*\pm}$  mesons were reconstructed using the “golden” decay channel  $D^{*\pm} \rightarrow D^0 \pi_s^\pm \rightarrow K^\mp \pi^\pm \pi_s^\pm$ , requiring  $|\eta(D^{*\pm})| < 1.5$  and  $P_T(D^{*\pm}) > 2 \text{ GeV}$ . Jets were reconstructed using the inclusive  $k_\perp$  algorithm, requiring  $|\eta_{\text{jet}}| < 2.4$  and  $E_{T,\text{jet}} > 9 \text{ GeV}$ . Since the jets were reconstructed as massless, the jet observable reduces to  $z_{\text{jet}} = (E + P_L)_{D^*} / 2E_{\text{jet}}$ . The contribution of  $D^{*\pm}$  mesons from B-hadron decays, which amounts to about 9%, was subtracted using the prediction of the PYTHIA Monte Carlo program.

The H1 collaboration measured the normalized  $D^{*\pm}$  meson cross sections as a function of both  $z_{\text{hem}}$  and  $z_{\text{jet}}$  in DIS ( $2 < Q^2 < 100 \text{ GeV}^2$  and  $0.05 < y < 0.7$ ), using the same decay channel and requiring  $|\eta(D^{*\pm})| < 1.5$  and  $1.5 < P_T(D^{*\pm}) < 15 \text{ GeV}$ . Jets were reconstructed using the massive inclusive  $k_\perp$  algorithm in the  $\gamma^* p$  rest-frame. The measurement was performed for two event samples. In the first sample, referred to as the “ $D^{*\pm}$  jet sample”, the presence of a jet containing the  $D^{*\pm}$  with  $E_{T,\text{jet}} > 3 \text{ GeV}$  is required as a hard scale. In the second sample, the “no  $D^{*\pm}$  jet sample”, no such jet is present. The small 1 – 2% contribution of  $D^{*\pm}$  mesons originating from B-hadron decays was estimated with the RAPGAP MC and was subtracted from the data. Both measurements were corrected for detector and QED radiative effects.

The corrected data, shown in figure 5, were used to fit the parameters of fragmentation functions for two classes of QCD models: 1) the leading-order + parton shower models as implemented in the Monte Carlo programs RAPGAP (used by H1) and PYTHIA (used by ZEUS), interfaced with the Lund string model for fragmentation as implemented in PYTHIA, and 2) the next-to-leading-order (NLO) QCD calculations as implemented in HVQDIS (fixed flavor number scheme) and used by H1 for DIS, and in FMNR (variable flavor number scheme) used by ZEUS for photoproduction, with charm quarks fragmented independently to  $D^{*\pm}$  mesons. For comparison of the data with NLO calculations, hadronization corrections have been applied.

The values of the fragmentation function parameters  $\varepsilon$  and  $\alpha$  extracted for the Peterson and Kartvelishvili parametrizations respectively can be found in table 1. The optimal (at  $\chi_{\text{min}}^2$ ) fragmentation parameter value depends on the settings used for other free parameters of the PYTHIA model. With the default settings, the parameters extracted by ZEUS and H1 for the  $D^{*\pm}$  jet sam-

ple are in good agreement. When using the PYTHIA parameter settings tuned by ALEPH [19] a harder fragmentation function is needed to describe the data. This can be understood as being due to a significant fraction of  $D^{*\pm}$  mesons produced in decays of higher excited charm states, provided by the ALEPH setting in contrast to the default setting. The resulting value of the Peterson parameter, extracted by H1 using the  $D^{*\pm}$  jet event sample, is in agreement with the value  $\varepsilon = 0.040$  extracted by ALEPH from their data. This result is also consistent with the hypothesis of fragmentation universality in  $ep$  and  $e^+e^-$  processes.

For H1, in case of the HVQDIS NLO calculation, the data are well described after fitting the Kartvelishvili parametrization, while when using the Peterson one no satisfactory description of the data is achieved. In the case of ZEUS, both parametrizations are able to describe the data.

FF parametrization	ZEUS: PHP	H1: DIS		
		$D^{*\pm}$ jet sample		No $D^{*\pm}$ jet sample
	$z_{\text{jet}}$	$z_{\text{jet}}$	$z_{\text{hem}}$	$z_{\text{hem}}$
PYTHIA with default parameter setting:				
Peterson ( $\varepsilon$ )	$0.064 \pm 0.06_{-0.008}^{+0.011}$	$0.061_{-0.009}^{+0.011}$	$0.049_{-0.010}^{+0.012}$	$0.010_{-0.002}^{+0.003}$
Kartvelishvili ( $\alpha$ )	—	$3.1_{-0.3}^{+0.3}$	$3.3_{-0.4}^{+0.4}$	$7.6_{-1.1}^{+1.3}$
PYTHIA with ALEPH parameter setting [19]:				
Peterson ( $\varepsilon$ )	—	$0.035_{-0.006}^{+0.007}$	$0.029_{-0.005}^{+0.007}$	$0.006_{-0.002}^{+0.002}$
Kartvelishvili ( $\alpha$ )	—	$4.3_{-0.4}^{+0.4}$	$4.5_{-0.5}^{+0.6}$	$10.3_{-1.6}^{+1.7}$
NLO calculations FMNR (PHP) and HVQDIS (DIS):				
Peterson ( $\varepsilon$ )	$0.0721_{-0.0123}^{+0.0139}$	$0.034_{-0.004}^{+0.004}$	$0.070_{-0.013}^{+0.015}$	$0.007_{-0.001}^{+0.001}$
Kartvelishvili ( $\alpha$ )	$2.87_{-0.35}^{+0.33}$	$3.8_{-0.3}^{+0.3}$	$3.3_{-0.4}^{+0.4}$	$6.0_{-0.8}^{+1.0}$

Table 1: Extracted fragmentation function parameters.

The hemisphere observable used by H1 allows to investigate charm fragmentation also close to the kinematic threshold, by selecting events which do not contain a  $D^{*\pm}$  jet above the minimal  $E_T$  cut. The corresponding normalised  $D^{*\pm}$  meson cross sections together with the prediction of RAPGAP with the Kartvelishvili FF fitted to the data are shown in figure 5 d. The extracted fragmentation function is found to be significantly harder than the one fitted to the  $D^{*\pm}$  jet sample (the dotted line). This can be interpreted as an inadequacy of the QCD model to provide a consistent description of the full phase space down to the kinematic threshold. The NLO HVQDIS calculation fails to describe this data sample.

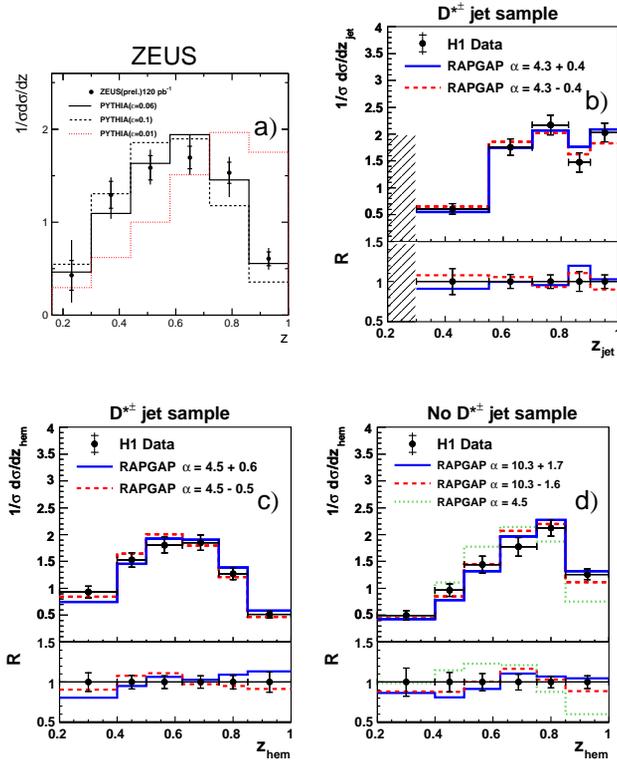


Fig. 5: Normalized  $D^{*\pm}$  meson cross sections as a function of the fragmentation observables: a)  $z_{\text{jet}}$  as measured by ZEUS, b)  $z_{\text{jet}}$  as measured by H1, c) and d)  $z_{\text{hem}}$  for the “ $D^{*\pm}$  jet” and the “no  $D^{*\pm}$  jet” samples. The full and dashed lines in the H1 sub-figures indicate a variation of  $\pm 1\sigma$  around the best fit value.

### 3 Beauty production at HERA

*Authors: S. Boutle, M. Turcato, A. Yagües-Molina*

#### 3.1 Introduction

At HERA, beauty quarks are produced predominantly via the boson-gluon fusion process, where a photon emitted by the electron interacts with a gluon in the proton producing a  $b\bar{b}$  pair. The measurements of such interactions are directly sensitive to the gluon density in the proton. Also, perturbative calculations of these processes should be reliable since the virtuality of the exchanged photon,  $Q^2$ , in the case of deep inelastic scattering (DIS), and the large mass of the produced quark, in the case of photoproduction, provide a hard scale. Hence, the study of  $b$  quark production at HERA is a stringent test of perturbative Quantum Chromodynamics (QCD). Measurements of such processes made at HERA are relevant for the LHC since they can test the precision of the description of  $b$  quark production by theoretical calculations. They also use tagging methods and event topologies which can be used to improve experimental techniques at the LHC. In the following, recent H1 and ZEUS measurements of beauty production are presented.

#### 3.2 Measurement of beauty photoproduction using semileptonic decays into leptons.

The installation of the silicon Micro-Vertex detector [20] (MVD) in the ZEUS detector [21] during the HERA luminosity upgrade period 2000/2001 allowed the heavy flavour measurements to reach higher precision. In one such measurement, beauty quarks were tagged by identifying a muon from the  $b$  semileptonic decay. The choice of a muon provides a clean experimental signature of the events. In this measurement two variables were used to discriminate between different quark decays. The first is the relative transverse momentum,  $p_T^{\text{rel}}$ , of the muon with respect to the heavy flavour hadron which for experimental purposes is approximated to the direction of the jet associated with the muon. This variable can be used to discriminate between beauty and charm decays since the mass of the beauty quark is larger, and therefore the  $p_T^{\text{rel}}$  spectrum for muons coming from  $b$  is harder. The second variable is the signed impact parameter,  $\delta$ , of the muon track. The absolute value of  $\delta$  is given by the transverse distance of closest approach of the track to the beam spot, where the beam spot position as a function of time is evaluated as the mean position of the event vertex over a proper event range. The sign of  $\delta$  is positive if the angle between the axis of the associated jet and the line joining the beam spot to the point of closest approach of the track is less than  $90^\circ$ , and is negative otherwise. The variable  $\delta$  reflects the lifetime of the quark and hence can be used to discriminate between charm and beauty decays and the decays of light quarks. The sign allows a statistical separation of detector resolution effects from the effects of the decay lifetime of the heavy hadron.

By fitting template distributions from Monte Carlo simulations of the  $p_T^{\text{rel}}$  and  $\delta$  variables to the data, the beauty fraction in the data can be extracted and used to calculate cross sections. The distributions of the  $p_T^{\text{rel}}$  and  $\delta$  variables are shown in Fig. 6 compared to MC predictions. The data are well described by the MC simulations.

The measurement presented here is based on a data sample collected during 2005 corresponding to an integrated luminosity of  $124 \text{ pb}^{-1}$ . Photoproduction ( $Q^2 < 1 \text{ GeV}^2$ ) events with  $0.2 < y < 0.8$ , having two jets with  $p_T^{j1,j2} > 7,6 \text{ GeV}$ ,  $|\eta^{j1,j2}| < 2.5$  and a muon with  $p_T^\mu > 2.5$

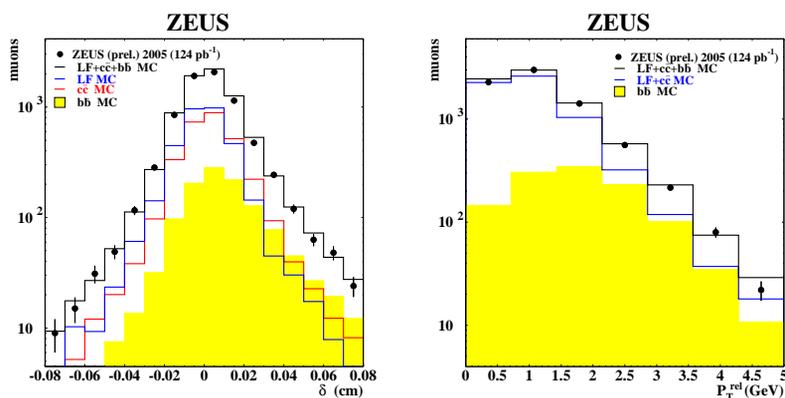


Fig. 6: Distribution of  $p_T^{\text{rel}}$  (left) and muon impact parameter  $\delta$  (right) of the data compared to the MC distributions for quarks of different flavour.

GeV and  $-1.6 < \eta^\mu < 2.3$  were selected. The event inelasticity,  $y$ , represents, in the proton rest frame, the fraction of the electron momentum which is transferred to the photon.

Figure 7 shows the distributions of the differential cross sections as a function of the muon transverse momentum,  $d\sigma/dp_T^\mu$ , and muon pseudorapidity  $d\sigma/d\eta^\mu$ . The results are compared to the ZEUS HERA-I data<sup>1</sup> [22] and to a NLO QCD prediction computed with the FMNR [14] program and corrected for hadronisation effects. The new results are in agreement with the previous measurement and compatible with NLO QCD predictions.

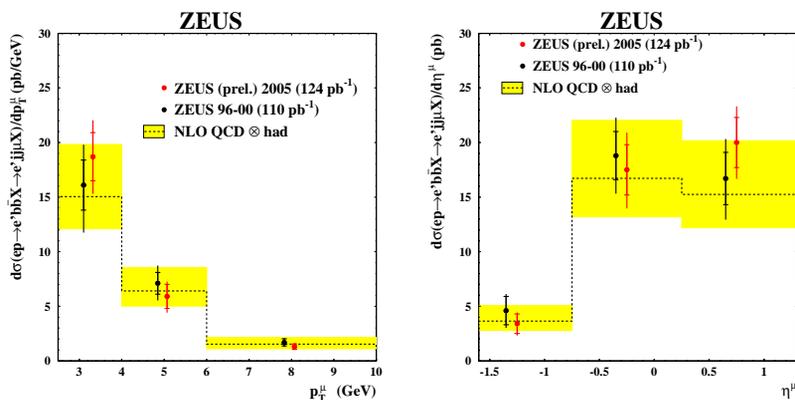


Fig. 7: Differential cross sections as a function of the muon transverse momentum,  $p_T^\mu$ , (left) and of the muon pseudorapidity,  $\eta^\mu$ , (right) for beauty photoproduction in dijet events with a muon. The measurements are compared to previous results and to NLO QCD predictions corrected for hadronisation effects.

Beauty photoproduction has been also measured using semileptonic decays to electrons or

<sup>1</sup>HERA-I refers to the data taken from 1996 to 2000 running period, previous to HERA luminosity upgrade.

positrons [23]. Tagging electrons has the advantage that lower values of the lepton transverse momentum are reachable. In this analysis, based on  $\mathcal{L} = 120 \text{ pb}^{-1}$  of HERA I data collected with the ZEUS detector from 1996 to 2000, events were selected in the photoproduction regime,  $Q^2 < 1 \text{ GeV}^2$ , having  $0.2 < y < 0.8$ , and with at least two jets with  $E_T^{j1,j2} > 7,6 \text{ GeV}$ ,  $|\eta^{j1,j2}| < 2.5$  and an electron coming from the semileptonic  $b$  decay with  $p_T^e > 0.9 \text{ GeV}$  and  $|\eta^e| < 1.5$ . For the identification of the electrons and the extraction of the  $b$  fraction a likelihood ratio method was used combining five discriminating variables. Three of them were used mainly for the lepton identification, and are based on the ionisation energy loss of the particle in the ZEUS central drift chamber, and on other calorimeter and tracking information. The other two are the momentum of the electron candidate transverse to the jet direction,  $p_T^{\text{rel}}$ , and the azimuthal angle between the electron and the missing transverse momentum vector, which corresponds to the neutrino from the semileptonic  $b$  decay. Figure 8 shows the distributions of the differential cross sections as a function of the electron transverse momentum,  $d\sigma/dp_T^e$ , and pseudorapidity,  $d\sigma/d\eta^e$ . The data are compared with the predictions of the PYTHIA MC program, scaled by a factor 1.75, and with NLO QCD predictions from FMNR. The shape of the data is well described by both the MC and the NLO calculations. The NLO predictions describe the normalisation of the data within the large uncertainties.

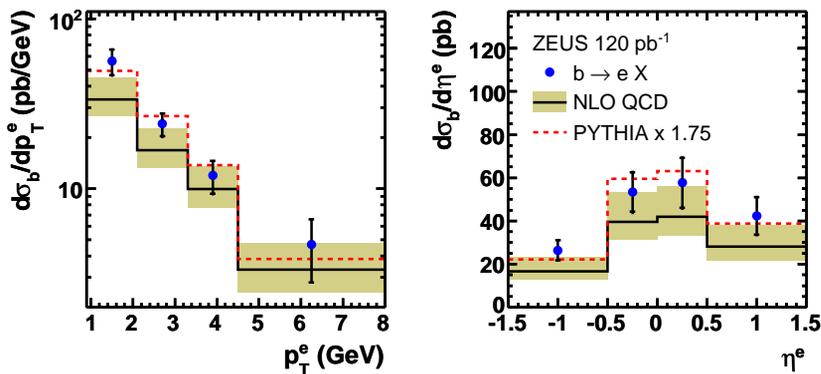


Fig. 8: Differential cross sections as a function of the transverse momentum,  $p_T^e$ , (left) and pseudorapidity,  $\eta^e$ , (right) of the electron for beauty photoproduction in dijet events with an electron. The measurements are compared to the predictions from PYTHIA as well as to NLO QCD calculations corrected for hadronisation effects.

### 3.3 Measurement of beauty dijet cross sections in photoproduction using inclusive lifetime tag.

An inclusive measurement of beauty in dijet events in the photoproduction regime [24] is presented here. The analysis is based on a sample of data collected by the H1 detector during the years 1999 and 2000 and corresponding to an integrated luminosity of  $56.8 \text{ pb}^{-1}$ . Photoproduction ( $Q^2 < 1 \text{ GeV}^2$ ) events with  $0.15 < y < 0.8$  and two jets with  $p_T^{j1,j2} > 11,8 \text{ GeV}$  and  $-0.9 < \eta^{j1,j2} < 1.3$  were selected.

Events containing beauty quarks were distinguished from those containing only light quarks by reconstructing the signed impact parameter,  $\delta$ , of the charged tracks, i.e. their distances to the primary vertex, using precise spatial information from the H1 vertex detector. The long lifetime of  $b$  flavoured hadrons lead to larger displacements than for light quark events.

The quantities  $S_1$  and  $S_2$  are defined as the significance,  $\delta/\sigma(\delta)$ , of the track with the highest and second highest absolute significance, respectively, where  $\sigma(\delta)$  is the error on  $\delta$ . In order to reject most of the light quark background and to reduce the uncertainty due to the impact parameter resolution, the negative bins in the significance distributions were subtracted from the positive ones. To extract the beauty fraction, a simultaneous  $\chi^2$ -fit to the subtracted  $S_1$  and  $S_2$  distributions was performed (see Fig. 9). The differential cross sections as a function of  $p_T^j$  and  $\eta^j$ , shown in Fig. 10, are extracted using the scale factors obtained from the fit. The results are compared to different MC predictions and to NLO QCD calculations. The beauty cross sections are reasonably well described in shape, whereas the NLO QCD prediction seems to lie below the data.

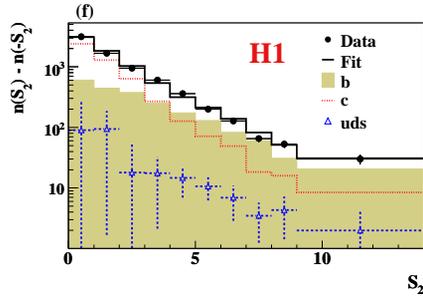


Fig. 9: Distributions of the subtracted signed significance for the sample with at least two tracks reconstructed in the Central Silicon Tracker.

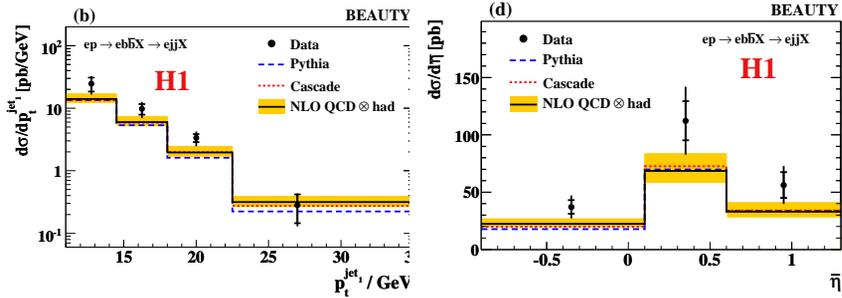


Fig. 10: Differential beauty dijet photoproduction cross sections as a function of the transverse momentum of the jet,  $d\sigma/dp_T^j$ , (left) and as a function of the jet pseudorapidity,  $d\sigma/d\eta^j$ , (right). The measurements are compared to the absolute predictions of PYTHIA and CASCADE as well as to NLO QCD calculations corrected for hadronisation effects.

### 3.4 Beauty production measurement using double tagging techniques

Beauty identification based on a single lepton tagging in dijet events is a powerful tool that allows to select a large event sample at HERA. However, the request of the presence of two jets in an event and the high background due to lighter flavour events does not allow the measurement of  $b$  quarks produced at very low transverse momenta, and therefore a total beauty cross section cannot be extracted. A way to access lower  $b$ -quark transverse momenta is to use double tagging techniques, by identifying two particles coming from the beauty decay. In this case, the cleaner event signature reduces significantly the background from non-beauty events.

An analysis [25] using this kind of approach identified beauty in events in which a  $D^*$  and a muon were found in the final state. Charm production is a background to this analysis, since a  $c\bar{c}$  pair in which one of the charm quarks hadronise into a  $D^*$  and the other produces a muon have a similar signature. However, in charm events the muon and the  $D^*$  lie in opposite hemispheres, while in the case of beauty production a muon and a  $D^*$  coming from the same  $B$  hadron lie in the same hemisphere, and in addition have opposite charges. Therefore, beauty and background can be separated by using the charge correlations and angular distributions of the muon with respect to the  $D^*$  meson.

The analysis uses a sample of  $\mathcal{L} = 114 \text{ pb}^{-1}$  of data corresponding to the full HERA I statistics collected by the ZEUS detector. The visible cross section was evaluated for unlike-sign  $D^*$ -muon events. This cross section was then extrapolated to the parton level and compared to NLO QCD predictions. No cut on the transverse momentum of the  $b$  quark,  $p_T^b$ , was imposed. The measured cross section in the kinematic region  $Q^2 < 1 \text{ GeV}^2$ ,  $\zeta^b < 1$ , where  $\zeta$  is the  $b$ -quark rapidity,  $0.05 < y < 0.85$  is

$$\sigma(ep \rightarrow b(\bar{b})X) = 11.9 \pm 2.9(\text{stat.})_{-3.3}^{+1.8}(\text{syst.}) \text{ nb}, \quad (1)$$

to be compared to a NLO QCD prediction of

$$\sigma^{NLO}(ep \rightarrow b(\bar{b})X) = 5.8_{-1.3}^{+2.1} \text{ nb}. \quad (2)$$

The measured cross section exceeds the NLO QCD prediction, but is compatible within the errors.

In another double-tagging analysis [26], events with two muons in the final state were used to study beauty production. This method has many advantages over the the  $D^*\mu$  analysis. It has larger statistics due to the higher branching ratio; the kinematic region is larger allowing the extraction of the total beauty cross section with almost no extrapolation; lower background induced by charm allows  $b\bar{b}$  correlations to be measured, testing the contribution of higher orders in perturbative calculations. The analysis uses  $114 \text{ pb}^{-1}$  of HERA I data collected by the ZEUS detector. The data sample is separated into high- and low-mass (isolated and non-isolated), like- and unlike-sign muon pairs. Since beauty is the only genuine source of like-sign muon pairs and fake muon background can give rise to like- and unlike-sign pairs, the beauty contribution can be determined from the difference between the like- and unlike-sign samples.

The kinematic region for the measurement of the total cross section was kept as large as possible:  $-2.2 < \eta^\mu < 2.5$ ,  $p_T^\mu > 1.5 \text{ GeV}$  for one muon and  $p_T^\mu > 0.75 \text{ GeV}$  for the other muon, as well as  $p > 1.8 \text{ GeV}$  for  $\eta < 0.6$ , or ( $p > 2.5 \text{ GeV}$  or  $p_T > 1.5 \text{ GeV}$ ) for  $\eta > 0.6$ .

Also in this case, a visible cross section was measured and then extrapolated to the total beauty cross section. DIS and photoproduction regimes were not separated. The measured total beauty cross section is

$$\sigma_{\text{tot}}(ep \rightarrow b\bar{b}X) = 13.9 \pm 1.5(\text{stat.})_{-4.3}^{+4.0}(\text{syst.}) \text{ nb.} \quad (3)$$

The NLO QCD prediction was obtained by adding the predictions from FMNR and HVQDIS [27] for the photoproduction and DIS parts, respectively:

$$\sigma_{\text{tot}}^{\text{NLO}}(ep \rightarrow b(\bar{b})X) = 7.5_{-2.1}^{+4.5} \text{ nb.} \quad (4)$$

Also in this case, the NLO QCD prediction is lower than the measured value, but compatible within the large uncertainties.

Visible differential cross sections were also measured, in the kinematic region defined by  $p_T^\mu > 1.5 \text{ GeV}$ ,  $-2.2 < \eta^\mu < 2.5$  for both the muons, in order to ensure a uniform kinematic acceptance. Figure 11 shows the the differential cross sections as a function of the muon transverse momentum,  $d\sigma/dp_T^\mu$ , and pseudorapidity,  $d\sigma/d\eta^\mu$ . The data are well described in shape by the theoretical predictions, with a tendency of the NLO QCD calculations to underestimate the normalisation of the data consistent with the observations from the total cross section.

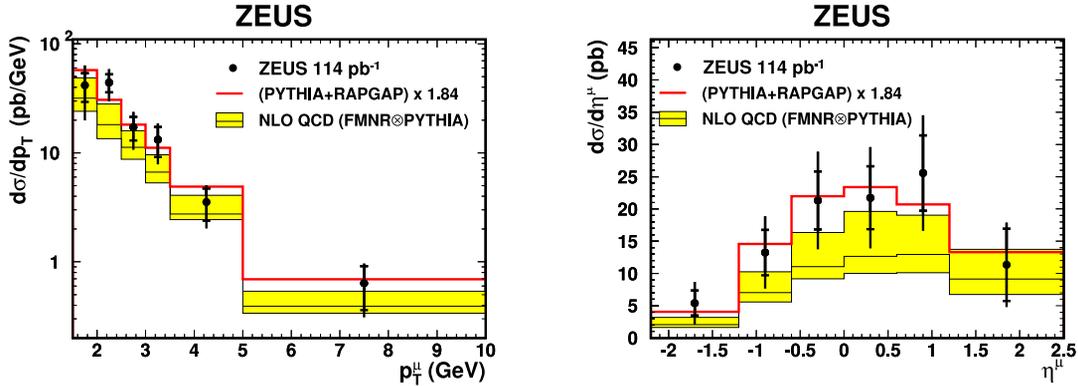


Fig. 11: Differential cross sections  $d\sigma/dp_T^\mu$  (left) and  $d\sigma/d\eta^\mu$  (right) for muons from  $b$  decays in dimuon events. The measurements (solid dots) are compared to the scaled sum of the predictions by the LO+PS generators PYTHIA and RAPGAP (histogram) and to the NLO QCD predictions from FMNR  $\otimes$  PYTHIA.

### 3.5 Conclusions

Beauty production at HERA is extensively studied using different analysis techniques. Beauty tagging with a single lepton gives a high statistics sample for the analyses, and the precision of the measurements is now comparable or better than that of the theoretical predictions. The measurements based on electron and muon tagging are affected by different systematic uncertainties, cover a slightly different kinematic region, and cross check each other. Inclusive analyses based on lifetime are also being done and will reach their full potential when the full HERAII data sample will be used. This kind of analyses are sensitive to beauty production also at large  $p_T^b$ .

For the investigation of the lower  $p_T^b$  region, double tagging techniques have been developed. In this way, the total beauty production cross section can be measured. Although these

measurements are still affected by a relatively large statistical uncertainty, they show that the difference between the observed cross sections and the theoretical predictions is not larger at lower transverse momenta.

The study of beauty production at HERA is significantly testing the precision of the perturbative QCD predictions, over a wide range in  $p_T$  and  $\eta$  of the produced  $b$  quarks. The understanding of beauty production in terms of perturbative QCD is vital for the future measurements that will be done at the large hadron collider, where a significant part of the cross section will consist on beauty.

## 4 Experimental Status of $F_2^{c\bar{c}}$ and $F_2^{b\bar{b}}$ at HERA

Authors: Philipp Roloff, Monica Turcato

### 4.1 Theoretical description

The double differential cross section versus  $x$  and  $Q^2$  for the production of a heavy quark (charm or beauty) pair,  $Q\bar{Q}$ , in deep inelastic scattering can be described by the heavy quark contributions to the proton structure functions:

$$\frac{d^2\sigma^{Q\bar{Q}}(x, Q^2)}{dx dQ^2} = \frac{2\pi\alpha^2}{Q^4 x} \left\{ [1 + (1-y)^2] F^{Q\bar{Q}}(x, Q^2) - y^2 F_L^{Q\bar{Q}}(x, Q^2) \right\}. \quad (5)$$

In the simplified picture of the quark-parton model (QPM) the electron scatters off a single quark in the proton. In this case  $x$  can be interpreted as the fraction of the proton momentum carried by the struck quark. Since heavy quarks can not exist within the proton due to their high mass, they are dominantly produced by the boson gluon fusion (BGF) process.

Heavy quark production as described above can be interpreted in two ways: on one hand it is possible to treat charm and beauty as massive quarks which are produced dynamically in the scattering process. In this case  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$  provide an indirect measurement of the gluon content of the proton. On the other hand it is possible to consider the splitting of a gluon into a heavy quark pair to happen within the proton for  $Q^2 \gg (2m_Q)^2$ . Hence here  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$  give the virtual charm and beauty content of the proton. As a consequence, the use of i.e.  $Z$  boson production as a luminosity monitor at the LHC requires a precise knowledge of the beauty content of the proton.

The large masses ( $m_c, m_b \gg \Lambda_{QCD}$ ) of the charm and beauty quarks provide an additional hard scale in perturbative QCD calculations. Different approaches exist to describe the multi scale problem of heavy quark production in  $ep$  collisions. In the massive or fixed flavour number scheme (FFNS) the proton contains only light quarks while charm and beauty are produced dynamically. Thus the threshold region is handled correctly, but the presence of other large scales, e.g.  $Q^2$  or the transverse momentum of the heavy quarks,  $p_T$ , can spoil the convergence of the perturbative expansion. In contrast, charm and beauty are treated as massless partons within the proton in the zero mass variable flavour number scheme (ZM-VFNS) which can improve the reliability of the calculations if one of the competing scales becomes large. An interpolation between both approaches is done in the (general mass) variable flavour number scheme (GM-VFNS) where heavy flavour production is treated as massive at low  $Q^2$  and massless at high  $Q^2$ .

While precise measurements of charm production are feasible using the large HERA II data sample, the measurements of beauty production are usually limited by the small production cross section. Since effects due to the higher beauty mass are relevant in a large part of the phase space accessible at HERA, beauty production might help to improve the understanding of mass effects for heavy quark production in deep inelastic scattering. A possible scenario is to “calibrate” theory predictions using beauty production and apply the improvements to charm for the extraction of the gluon content of the proton.

## 4.2 Experimental results on $F_2^{c\bar{c}}$

The charm contribution to the proton structure function  $F_2$ ,  $F_2^{c\bar{c}}$ , has been measured at HERA by the two Collaborations ZEUS [28–31] and H1 [32–35], in a wide kinematic region in  $x$  ( $0.00002 \lesssim x \lesssim 0.03$ ) and in the photon virtuality,  $Q^2$  ( $1 < Q^2 < 1000 \text{ GeV}^2$ ).

Charm production at HERA can be tagged in different ways. In the so-called *golden mode* a  $D^*(2010)$  meson is reconstructed through its decay  $D^{*+} \rightarrow D^0 \pi^+ \rightarrow K^- \pi^+ \pi^+$  (+c.c.). Other charmed mesons can also be reconstructed: the most copiously produced are  $D^0$ ,  $D^\pm$ ,  $D_s$ . The production cross sections of all these mesons can be measured in a defined kinematic region, and the total charm cross section can then be extracted by extrapolating the measurements to the full phase space. This extrapolated cross section is then used to evaluate  $F_2^{c\bar{c}}$ . Both the ZEUS and the H1 Collaborations have used this method to extract  $F_2^{c\bar{c}}$  from  $D$  mesons cross sections. Charm tagging with mesons gives a clean signature of charm production, but the extrapolation to the total charm cross section can be large, especially in the low  $Q^2$  region (as an example, at ZEUS typical extrapolation factors range from  $\sim 4$  in the low- $Q^2$ , low- $x$  region to 1.5 at high- $Q^2$  [28]).

An alternative method to tag charm production takes advantage of the long lifetime of the charmed particles, by reconstructing secondary vertices from  $D$ -meson decay products, or, in inclusive analyses, by identifying tracks having impact parameter,  $\delta$ , significantly displaced from the event vertex. The secondary vertex reconstruction for  $D^\pm$  and  $D^0$  mesons has been used by the ZEUS Collaboration to enhance the signal to background ratio, and therefore the statistical precision, of the measurement [31]. On the other side, fully inclusive analyses use the significance of the impact parameter of the highest impact parameter tracks to separate charm and beauty from light flavour production [33, 34], since heavy flavours show a longer tail in the positive side of this distribution. The fraction of charm and beauty in an inclusive data sample can therefore be extracted by fitting the significance distribution to the contributions from beauty, charm, and lighter quarks. This method has been used by the H1 Collaboration to obtain some of the results presented here [33–35].

The advantage of the inclusive method is that the kinematic region for the measurement of charm production is significantly enlarged, and therefore the extrapolation needed for the measurement of  $F_2^{c\bar{c}}$  is strongly reduced.

The program HVQDIS [27] is the only program which is able to provide theoretical predictions for  $D$  meson production cross sections at NLO accuracy in perturbative QCD. It was used by the ZEUS and H1 Collaborations to extrapolate the measured cross section for a particular  $D$  meson final state to  $F_2^{c\bar{c}}$ . In this program, the production of heavy flavours is performed using the fixed flavour number scheme. The ZEUS and H1 measurements of  $F_2^{c\bar{c}}$  extracted in this way should therefore be compared with NLO QCD predictions evaluated in the FFNS.

The results for  $F_2^{c\bar{c}}(x, Q^2)$  are shown in Fig. 12. In the figure the ZEUS and H1 measurements, obtained from charmed meson production, are compared with the H1 results from inclusive lifetime measurements. The agreement between the experiments is good, validating the two different analysis procedures. The data rise with increasing  $Q^2$ , with the rise becoming steeper at lower  $x$ .

The data are also compared with perturbative QCD predictions at NLO. Two different parameterisations of the proton PDFs have been used for the NLO QCD calculations, in or-

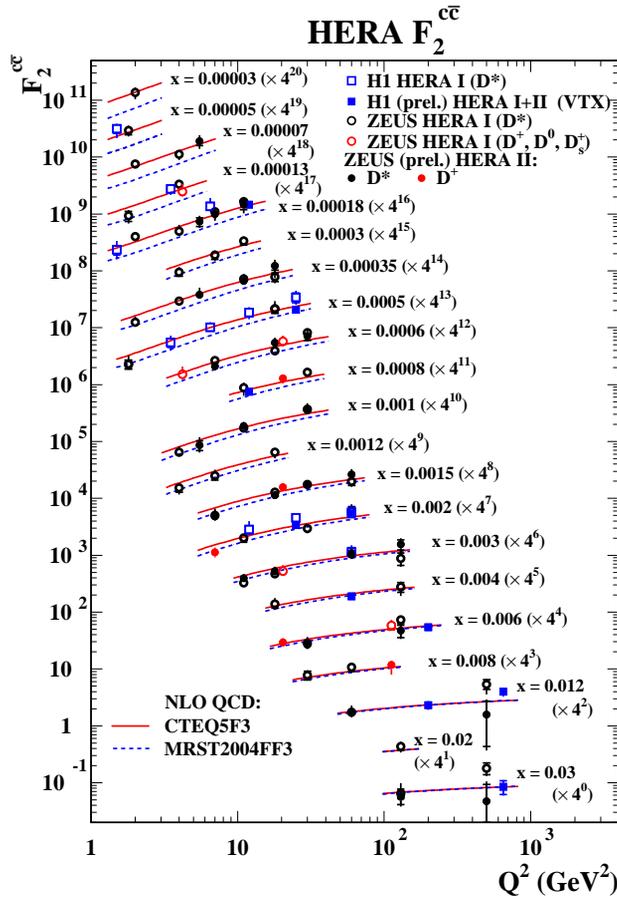


Fig. 12: The measured  $F_2^{c\bar{c}}$  at  $x$  values between 0.00003 and 0.03 as a function of  $Q^2$ . The data are shown with statistical uncertainties (inner bars) and statistical and systematic uncertainties added in quadrature (outer bars). The data are compared with next-to-leading order QCD predictions, evaluated using different proton PDFs.

der to check the sensitivity of the predictions to different gluon densities: CTEQ5F3 [36] and MRST2004FF3 [37]. The charm data are in general well described by NLO QCD: this shows that the proton PDFs, which are extracted mainly from inclusive scattering data, are also able to describe reasonably well charm production. The two PDFs show differences in the low  $x$  region, demonstrating the sensitivity of the measurement to different parameterisation of the gluon density in the proton.

### 4.3 Experimental results on $F_2^{b\bar{b}}$

The beauty contribution to the inclusive structure function  $F_2$  was measured by the ZEUS Collaboration using muons and jets and by the H1 Collaboration from lifetime information of displaced tracks.

The ZEUS Collaboration measured beauty production in events with a muon and a jet,

using a data sample of  $\mathcal{L} = 39 \text{ pb}^{-1}$ . The fraction of beauty quarks in the data was derived using the distribution of the transverse momentum of the muon relative to the axis of the associated jet,  $p_T^{\text{rel}}$  [38]. Due to the larger mass of the beauty quark, muons originating from  $b$  decays tend to higher values of  $p_T^{\text{rel}}$  compared to muons from charm and light flavour decays. The beauty contribution to  $F_2$ ,  $F_2^{bb}$ , was obtained by extrapolating the double differential cross sections as a function of  $Q^2$  and  $x$  to the full phase space using the HVQDIS program, as for the extraction of  $F_2^{c\bar{c}}$  from the visible cross sections for  $D$  meson production. Here extrapolation factors between 3 and 6 decreasing with  $Q^2$  had to be applied.

The H1 collaboration extracted  $F_2^{bb}$  in a fully inclusive analysis based on information from the Central Silicon Tracker. The impact parameter significance of tracks in the transverse plane was used in a fit to extract the (charm and) beauty fractions in the considered data sample [35]. Due to the long lifetime of the  $B$  hadrons, it is possible to distinguish the position of the decay vertices of these particles from the primary interaction vertex. As a consequence, tracks originating from beauty decays exhibit large positive impact parameters compared to tracks coming from lighter quarks. An advantage of the inclusive lifetime method is that the extrapolation to the full phase space is smaller. Recent results from the HERA II period were combined with earlier measurements [33, 34].

The results obtained by the H1 and ZEUS Collaborations are summarised in Fig. 13. The reduced cross section

$$\tilde{\sigma}^{b\bar{b}} = F_2^{b\bar{b}} - \frac{y^2}{1 + (1 - y)^2} F_L^{b\bar{b}}, \quad (6)$$

is shown as a function of  $x$  for different values of  $Q^2$ . Although very different methods have been used, the results are in agreement within the large errors.

The data are compared to NLO QCD predictions using different schemes [39]. The CTEQ5F4 [36] is done in the FFNS, while MRST04 [40], MRST NNLO [41] and CTEQ6.5 [42] implement the VFNS. At low values of  $Q^2$  and  $x$  the predictions of the CTEQ and MRST groups differ up by a factor two, but the present statistical accuracy of the data does not allow to discriminate between the different calculations. A better precision of the data is needed in order to better understand the different aspects of the theoretical calculations and to disentangle between different approaches.

#### 4.4 Conclusions and outlook

Both experiments, H1 and ZEUS, collected a data sample of about  $0.5 \text{ fb}^{-1}$ . The analysis of the full HERA I+II dataset will increase the available statistics by a factor of 2 to 10, depending on the analysis. The combination of different heavy flavour tagging methods (e.g. different  $D$  mesons for charm, different leptons for beauty, inclusive analyses) can further improve the precision of the measurements, keeping also into account the fact that the systematic uncertainties of different tagging techniques are at least partially uncorrelated. The final step is the combination of the ZEUS and H1 data into a single measurement: this will again double the available dataset.

New detector components, which allow to extend the kinematic range of the  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$  measurements, were installed for the HERA II data taking period. The forward region can be studied using the ZEUS Straw Tube Tracker and Forward Microvertex Detector while the H1

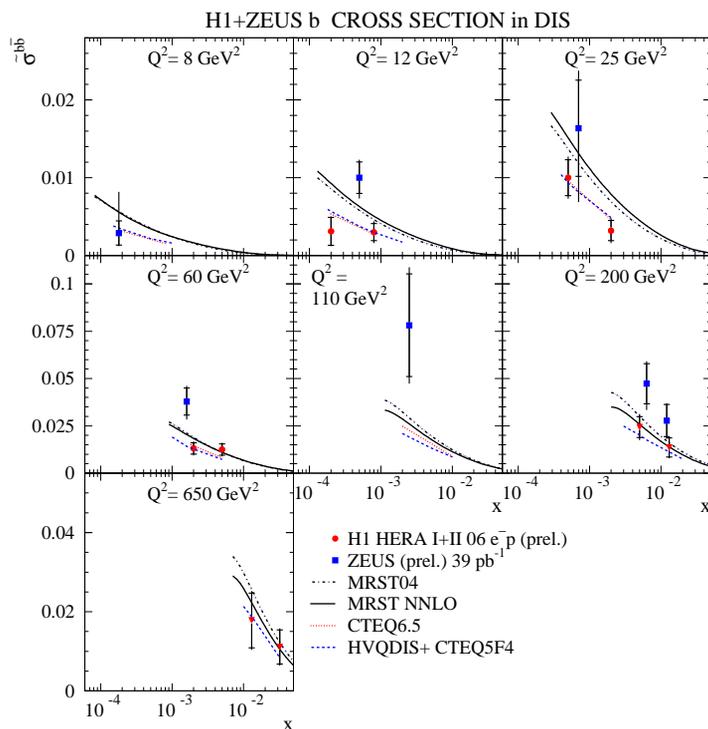


Fig. 13:  $\bar{\sigma}^{b\bar{b}}$  as a function of  $x$  for different values of  $Q^2$ . The inner error bars represent the statistical uncertainties only while the outer error bars correspond to statistical and systematic uncertainties added in quadrature. Different predictions are compared to the data.

Backward Silicon Tracker gives access to the backward region.

Significant improvement in the precision of the measurements is therefore possible and will be reached by the final HERA analyses. The new beauty measurements will be of help to understand some aspects of the theory that will be then implemented for the description of charm production. A final combined ZEUS+H1 measurement of charm production could be used in the PDF fits and will hopefully help to learn something more on the gluon density in the proton.

## References

- [1] H1 Collaboration, H1prelim-08-072,  
[https://www-h1.desy.de/publications/H1preliminary.short\\_list.html#HQ](https://www-h1.desy.de/publications/H1preliminary.short_list.html#HQ), 2008.
- [2] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. C **51**, 271 (2007).
- [3] H1 Collaboration, H1prelim-08-073,  
[https://www-h1.desy.de/publications/H1preliminary.short\\_list.html#HQ](https://www-h1.desy.de/publications/H1preliminary.short_list.html#HQ) (2008).
- [4] H1 Collaboration, A. Aktas *et al.*, arXiv 0808.1003 (hep-ex) 2008.

- [5] ZEUS Collaboration, ZEUS-prel-07-008,  
[http://www-zeus.desy.de/physics/hfla/public/abstracts07/paper/f2charm\\_zeus\\_EPSPaper\\_106.ps](http://www-zeus.desy.de/physics/hfla/public/abstracts07/paper/f2charm_zeus_EPSPaper_106.ps),  
 2008.
- [6] ZEUS Collaboration, S. Chekanov et al., Phys. Lett. **B649**, 111-121 (2007).
- [7] ZEUS Collaboration, S. Chekanov et al., DESY-08-129, arXiv:0807.1290 [hep-ex] (2008).
- [8] ZEUS Collaboration, S. Chekanov et al., JHEP **0707**, 74 (2007).
- [9] ZEUS Collaboration, ZEUS-prel-07-010,  
<http://www-zeus.desy.de/physics/hfla/public/abstracts07/>, 2008.
- [10] A. Baird et al., IEEE Trans. Nucl. Sci. **48**, 1276 (2001).
- [11] A. Schöning, Nucl. Instr. and Meth. **A518**, 542 (2004).
- [12] A.W. Jung et al., *First Results from the Third Level of the H1 Fast Track Trigger*, Proc. of the 2007 IEEE NPSS RealTime Conference, Chicago, USA, 2007.
- [13] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C 50**, 251 (2006).
- [14] S. Frixione, M.L. Mangano, P. Nason and G. Ridolfi, Phys. Lett. **B 348**, 633 (1995).
- [15] U. Bassler, G. Bernardi, Nucl. Instr. and Meth. **A361**, 197 (1995).
- [16] H. Jung, PDF4MC, <http://indico.cern.ch/conferenceDisplay.py?confId=36364>, 2008.
- [17] H1 Collaboration, A Aktas *et al.*, *Study of Charm Fragmentation into  $D^{*\pm}$  Mesons in Deep-Inelastic Scattering at HERA*, [arXiv:0808.1003], submitted to Eur. Phys. J. C
- [18] S. Fang [ZEUS Collaboration], “Charm fragmentation function and charm fragmentation fractions at ZEUS,” Proceedings DIS 2007, Munich, Germany, April 16 – 20, Eds. G. Grindhammer and K. Sachs, 10.3360/dis.2007.147
- [19] ALEPH Collaboration, S. Schael *et al.*, Phys. Lett. **B 606**, 265 (2005); G. Rudolph [ALEPH collaboration], private communication.
- [20] E.N. Koffeman *et al.*, Nucl. Instrum. Meth. **A 453**, 89 (2000).  
 D. Dannheim *et al.*, Nucl. Instrum. Meth. **A 505**, 663 (2003).
- [21] ZEUS Collaboration, U. Holm(ed.), *The ZEUS Detector*. Status Report (Unpublished), DESY (1993), available on  
<http://www-zeus.desy.de/bluebook/bluebook.html>
- [22] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Rev. **D 70**, 012008 (2004).
- [23] ZEUS Collaboration, S. Chekanov *et al.*, DESY-08-056 (May 2008).
- [24] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C47**, 597-610 (2006).

- [25] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C50** 299-314 (2007).
- [26] ZEUS Collaboration, S. Chekanov *et al.*, DESY-08-129 (2008).
- [27] B.W. Harris and J. Smith, Phys. Rev. **D 57**, 2806 (1998).
- [28] ZEUS Collaboration, S. Chekanov *et al.*, Phys. Rev. **D 69**, 012004 (2004).
- [29] ZEUS Collaboration, S. Chekanov *et al.*, JHEP **07**, 074 (2007).
- [30] M. Turcato [ZEUS Collaboration], Proc. of the 2007 Europhysics Conference on High Energy Physics, Manchester, England, July 2007,  
[http://www.iop.org/EJ/article/1742-6596/110/2/022054/jpconf8\\_110\\_022054.pdf](http://www.iop.org/EJ/article/1742-6596/110/2/022054/jpconf8_110_022054.pdf)
- [31] D. Nicholass [ZEUS Collaboration], Proc. of 15th Int. Workshop on Deep-Inelastic Scattering and Related Subjects, Munich, April 2007, <http://dx.doi.org/10.3360/dis.2007.145>
- [32] H1 Collaboration, A. Aktas *et al.*, Phys. Lett. **B 528**, 199 (2002).
- [33] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C 45**, 23 (2006).
- [34] H1 Collaboration, A. Aktas *et al.*, Eur. Phys. J. **C 40**, 349 (2005).
- [35] H1 Collaboration, *Submitted to the 23rd International Symposium on Lepton-Proton Interactions at High Energy, LP2007*, Daejeon, Republic of Korea, August 13-18, 2007.
- [36] CTEQ Collaboration, Eur. Phys. J. **C 12**, 375 (2000).
- [37] A. D. Martin, W. J. Stirling and R. S. Thorne, Phys. Lett. **B 636**, 259 (2006).
- [38] B. Kahle [ZEUS Collaboration], Proc. of 15th Int. Workshop on Deep-Inelastic Scattering and Related Subjects, Munich, April 2007, <http://dx.doi.org/10.3360/dis.2007.164>
- [39] P.D. Thompson, arXiv:hep-ph/0703103v1, 2007.
- [40] A.D. Martin *et al.*, Eur. Phys. J. **D 73**, 050419 (2006).
- [41] R.S. Thorne, Phys. Rev. D **73**, 050419 (2006).
- [42] W.K. Tung *et al.*, JHEP **02**, 259 (2006).

# Experimental study of heavy flavour production at RHIC and LHC

*M. Biasini<sup>a</sup>, C. Bombinati<sup>b</sup>, G.E. Bruno<sup>c</sup>, E. Lytken<sup>d</sup>, A. Mischke<sup>e</sup>, C. Rosemann<sup>f</sup>,  
A. Starodumov<sup>g\*</sup>, D. Stocco<sup>h</sup>, R. Wolf<sup>f</sup>, and M. zur Nedden<sup>i</sup>*

<sup>a</sup> University and INFN, Perugia, Italy

<sup>b</sup> University and INFN, Padua, Italy

<sup>c</sup> University and INFN, Bari, Italy

<sup>d</sup> CERN, Geneva, Switzerland

<sup>e</sup> Institute for Subatomic Physics, Faculty of Science, Utrecht University, Utrecht, the Netherlands

<sup>f</sup> Inst. für Experimentalphysik, Universität Hamburg and DESY, Germany

<sup>g</sup> Paul Scherrer Institut, PSI Villigen, Switzerland

<sup>h</sup> University and INFN, Torino, Italy

<sup>i</sup> Humboldt-University of Berlin, Germany

## Abstract

After reviewing the main heavy flavour results from experiments at the Relativistic Heavy Ion Collider (RHIC), we present the expected performance for some of the most significant measurements in the heavy flavour sector at the Large Hadron Collider (LHC), for the experiments ALICE, ATLAS, and CMS.

*Coordinator: A. Dainese*

## 1 Heavy flavour physics at RHIC

*Author: A. Mischke*

### 1.1 Introduction

Measurements at RHIC have revealed strong modification of the jet structure in high-energy nuclear collisions due to the interaction of hard scattered partons with the hot and dense medium created in these reactions. The study of heavy-quark (charm and bottom) production in the medium offers unique opportunities for the investigation of the properties of the Quark-Gluon Plasma (QGP). Heavy quarks are believed to be produced predominantly in hard scattering processes in the early stage of the collision, and they probe the produced medium as they propagate through it [1]. Due to their higher mass, the penetrating power is much higher for heavy quarks than for light quarks, providing a sensitive probe of the medium. The energy loss of heavy quarks in the medium is expected to be smaller compared to light quarks due to the mass dependent suppression of the gluon radiation under small angle, known as the dead-cone effect [2, 3].

---

On leave from ITEP, Moscow, Russia

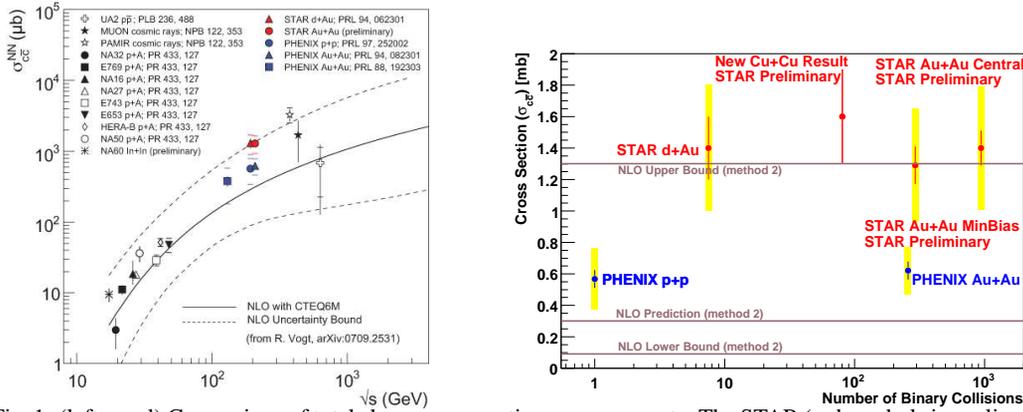


Fig. 1: (left panel) Comparison of total charm cross section measurements. The STAR (red symbols in on-line version) and PHENIX results (blue symbols in on-line version) are given as cross section per binary collisions. The dashed curves depict the uncertainty bands of the NLO calculations. (right panel) Total charm cross section divided by the number of binary collisions systems, compared to NLO calculations (horizontal lines).

## 1.2 Total charm cross section

The total charm cross section is currently determined through basically three different measurements: direct reconstruction of  $D$  mesons, muons and electrons. Electron identification in the PHENIX experiment is based on the Ring Imaging Cherenkov detector (RICH) in conjunction with a highly granular calorimeter. The particle momentum is measured by drift and pad chambers. The subtraction of the electron background (mainly from photons,  $\pi^0$  and  $\eta$ ) is performed by the converter and the cocktail methods [4, 5], which give similar results. In the STAR experiment, electrons are identified using the  $dE/dx$  and momentum measurements from the TPC together with the Time of Flight (ToF) information at low  $p_T$  ( $< 4 - 5$  GeV/c) and energy ( $E$ ) and shower shape measurement in the electromagnetic calorimeter (EMC) at high- $p_T$  ( $> 1.5$  GeV/c). The background contribution to the electrons from photonic sources are subtracted statistically [6].

The total charm cross section is extracted from a combined fit to the measured particle spectra. The STAR data are from combined fits to hadronic and semileptonic decay data. The PHENIX data are from semileptonic decay measurements only. The total cross sections from STAR and PHENIX are compared to results at other energies and to NLO calculations [7] in Fig. 1 (left panel). The discrepancy between STAR and PHENIX is under investigation. The data agree with the NLO prediction on the total charm cross section. The large theoretical uncertainty leads to a little predictive power in the total charm cross section. Fig. 1 (right panel) depicts the charm cross section divided by the number of binary collisions for different collisions systems. Within errors, the charm cross section for the different collisions systems follows binary collisions scaling, supporting the assumption that charm is predominantly produced by hard scattering in the initial state of the collision.

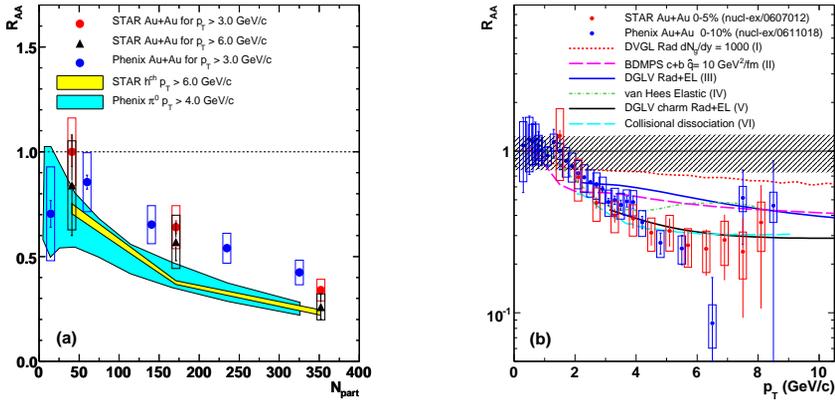


Fig. 2: Nuclear modification factor  $R_{AA}$  of non-photonic electrons in Au+Au collisions. (a)  $R_{AA}$  as a function of  $N_{part}$ . (b)  $R_{AA}$  as a function of  $p_T$  for the most central collisions.

### 1.3 Heavy-quark energy loss in hot and dense QCD matter

Nuclear effects are typically quantified using the nuclear modification factor  $R_{AA}$  where the particle yield in Au+Au collisions is divided by the yield in  $pp$  reactions scaled by the number of binary collisions.  $R_{AA} = 1$  would indicate that no nuclear effects, such as Cronin effect, shadowing or gluon saturation, are present and that nucleus-nucleus collisions can be considered as a incoherent superposition of nucleon-nucleon interactions. The average  $R_{AA}$  for high- $p_T$  non-photonic electrons as a function of participating nucleons ( $N_{part}$ ) is illustrated in Fig. 2(a). The STAR and PHENIX  $R_{AA}$  for non-photonic electrons are consistent with each other and shows an increasing suppression from peripheral to central Au+Au collisions, indicating an unexpectedly energy loss of heavy quarks in the medium in contradiction to expectations from the dead-cone effect. The suppression is similar to the one observed for light-quark hadrons, indicated by the shaded area in the figure. Fig. 2(b) shows the  $p_T$  dependence  $R_{AA}$  of non-photonic electrons in central Au+Au collisions. A strong suppression of a factor of  $\sim 5$  is observed for  $p_T > 6$  GeV/c. The  $R_{AA}$  is compared to several theoretical model calculations [4, 6]. The observed suppression is overpredicted by the models using reasonable model parameters. The data is described reasonably well if the bottom contribution to the electrons is assumed to be small. Therefore, the observed discrepancy could indicate that the  $B$  dominance over  $D$  mesons starts at higher  $p_T$ . A possible scenario for  $B$  meson suppression invokes collisional dissociation in the medium.

### 1.4 Heavy-quark azimuthal correlations

The measurement of the relative charm and bottom contributions to the non-photonic electrons is essential for the interpretation of the non-photonic electron spectra. Azimuthal angular correlations between non-photonic electrons and hadrons allow to identify the underlying production process [8]. Heavy flavours have, in general, a harder fragmentation function than gluons and light quarks, making the near-side correlation more sensitive to the decay kinematics. For the same electron transverse momentum the near-side  $e$ -hadron angular correlation from  $B$  decays is much broader than that from  $D$  decays. Fig. 3 (left panel) shows the azimuthal correlation

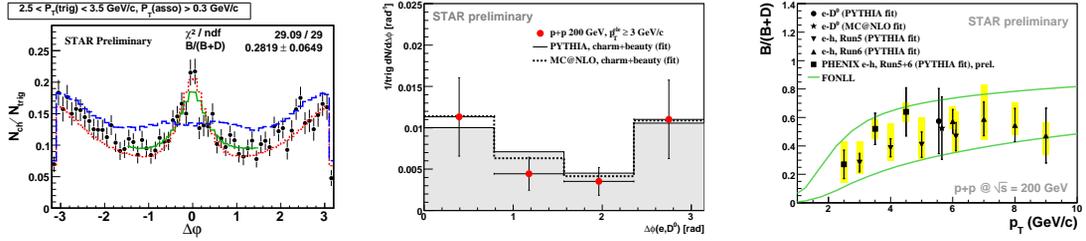


Fig. 3: (left panel)  $e$ -hadron azimuthal correlation distribution in 200 GeV  $pp$  collisions, compared to PYTHIA simulations (colored curves). (middle panel)  $e - D^0$  azimuthal correlation distribution for like-sign  $e - K$  pairs. The grey histogram (dashed line) illustrates results from PYTHIA (MC@NLO) simulations. (right panel) Relative bottom contribution to the total non-photonic electron yield derived from  $e - D^0$  and  $e$ -hadron correlations, compared to the uncertainty band from a FONLL calculation.

function of non-photonic electrons and hadrons in  $pp$  collisions at  $\sqrt{s} = 200 \text{ GeV}$ . The data is fitted with a linear combination of the simulated charm and bottom distribution, obtained from PYTHIA simulations, to extract the relative bottom contribution  $B/(B + D)$ . Similar studies are performed for  $e - D^0$  azimuthal correlations [8] (cf. Fig. 3, middle panel). Moreover, it has been shown that higher order sub-processes like gluon splitting may have a significant contribution to the near-side correlation. This contribution is studied by indentifying the  $D^*$  content of jets [8]. The results indicate that gluon splitting to  $c\bar{c}$  pairs contributes about 5% of the open charm production observed at RHIC, consistent with predictions from MC@NLO calculations [9].

The relative bottom contribution  $B/(B + D)$  is shown in Fig. 3 (right panel) together with predictions from FONLL calculations [10]. These data provide convincing evidence that bottom contributes significantly ( $\sim 50\%$ ) to the non-photonic electron yields at high- $p_T$ . Further studies have to show whether these findings imply substantial energy loss of bottom quarks in the produced medium.

## 1.5 Summary and conclusions

The measured total charm cross section follows binary collisions scaling as expected from the assumption that charm is produced exclusively in initial hard scattering processes. The high- $p_T$  suppression of the non-photonic electron yield in Au+Au collisions is much larger than expected. Theoretical explanations are yet inconclusive. The bottom contribution in the non-photonic electron spectrum is studied by  $e$ -hadron and  $e - D^0$  correlations. First measurements on the charm content in jets shows that the gluon splitting contribution is small at RHIC. Detailed and systematic studies will be possible with heavy flavour measurements in the ALICE experiment at the CERN-LHC [11].

### Acknowledgments

The author thanks the organizers for the stimulating atmosphere during the workshop. This work is supported by a Veni grant from the Netherlands Organization for Scientific Research (project number: 680-47-109).

## 2 Heavy flavour program of the ALICE experiment at the LHC

*Authors: C. Bombonati, G.E. Bruno, and D. Stocco*

### 2.1 Introduction

The Large Hadron Collider (LHC) will produce proton–proton, Pb–Pb, other lighter systems like Ar–Ar and proton induced nucleus collisions up to the energies corresponding to the maximum magnetic rigidity of 23,350 Tm (e.g., proton–proton and Pb–Pb collisions at centre-of-mass energy per nucleon–nucleon  $\sqrt{s_{NN}} = 14$  TeV and 5.5 TeV, respectively). ALICE [12, 13] is the dedicated heavy-ion experiment at the LHC; its main physics goal is the study of strongly-interacting matter in the conditions of high-energy density ( $> 10$  GeVfm $^{-3}$ ) and high temperature ( $\gtrsim 0.2$  GeV) over large volume ( $10^2$ – $10^3$  fm $^3$ ), expected to be reached in central Pb–Pb collisions. The ALICE apparatus [12, 13] has excellent capabilities for heavy-flavour measurements, for both open heavy-flavoured hadrons and quarkonia. In this paper, we shall limit the discussion to the detection of open charm and beauty in the central barrel (section 2.2) and of quarkonium states at forward rapidity (section 2.4), with an emphasis on the proton–proton collisions. Therefore only the detectors involved in these analyses are described in the following.

The ALICE central barrel covers the pseudo-rapidity region  $-0.9 < \eta < 0.9$  and is equipped with tracking detectors and particle identification systems embedded in a magnetic field  $B = 0.5$  T. The combined information from the central barrel detectors allows to track charged particles down to low transverse momenta (low  $p_T$  cut-off  $\approx 100$  MeV/ $c$ ) and provides hadron and electron identification as well as an accurate measurement of the positions of the primary (interaction) vertex and of the secondary (decay) vertices. The main tracking detector is the Time Projection Chamber (TPC) which provides track reconstruction and particle identification via  $\frac{dE}{dx}$ . The Inner Tracking System (ITS) is the innermost central barrel detector and is composed of six cylindrical layers of silicon detectors. The two layers closest to the beam pipe (at radii of  $\approx 4$  and 7 cm) are equipped with pixel detectors, the two intermediate layers (radii  $\approx 15$  and 24 cm) are made of drift detectors, while strip detectors are used for the two outermost layers (radii  $\approx 39$  and 44 cm). The ITS is a key detector for open heavy-flavour studies because it allows to measure the track impact parameter (i.e. the distance of closest approach of the track to the primary vertex) with a resolution better than 50  $\mu\text{m}$  for  $p_T > 1.3$  GeV/ $c$ , thus providing the capability to detect the secondary vertices originating from heavy-flavour decays. Two other systems play an important role in the heavy-flavour analyses as far as particle identification is concerned. They are the Transition Radiation Detector (TRD) for high-momentum electron identification and the Time-Of-Flight (TOF) for pion, kaon and proton separation. All these four detectors have full azimuthal coverage.

The detection of heavy quarkonia in the di-muonic decay channel is performed by the ALICE Muon Spectrometer in the forward pseudo-rapidity region  $2.5 < \eta < 4$ . The detector consists of five tracking stations with two planes of Multi-Wire Proportional Chambers each, with a spatial resolution of about 100  $\mu\text{m}$ , a dipole magnet with an integral field of 3 Tm and two trigger stations of Resistive Plate Chambers placed behind an iron-wall muon filter with a thickness of about 7 interaction lengths. The system is completed by a front absorber of composite material, predominantly made of carbon and concrete, which is placed at 90 cm from the

interaction vertex to reduce the free decay length of pions and kaons, and a beam shield made of tungsten, lead and stainless steel to protect the chambers from particles and secondaries produced at large rapidities. The spectrometer can detect quarkonia down to  $p_T = 0$  and is designed to achieve an invariant-mass resolution of 70 (100) MeV/ $c^2$  at 3 (10) GeV/ $c^2$ , needed to resolve the  $J/\psi$  ( $\Upsilon$ ) resonances.

## 2.2 Open heavy flavour in the ALICE Central Barrel

Heavy flavours are produced in initial parton-parton interactions, in the early stage of the collision. Their production can be calculated to a reasonable degree of precision within pQCD and they offer the possibility to explore the properties of the medium created in the collision with probes of known mass and colour charge. The energy loss by gluon radiation, for instance, is expected to be parton-specific (stronger for gluons than for quarks due to the larger colour charge of gluons) and flavour-specific (stronger for lighter than for heavier quarks, due to the dead cone effect [14, 15]). In addition, the measurement of open heavy-flavour production is of an essential practical interest for quarkonium physics as well, both as a natural reference and B meson decays being a sizable source of non-prompt  $J/\psi$  in high energy collisions.

In figure 4 we compare schematically the ALICE  $p_T$  vs.  $\eta$  acceptance for charm (c) and beauty (b) hadrons to that of the other LHC experiments, for proton–proton collisions at  $\sqrt{s} = 14$  TeV. In this plot the high  $p_T$  reach is the one expected for one year of running at nominal luminosity (note that the value of the luminosity is different for each experiment:  $10^{34}$  cm $^{-2}$ s $^{-1}$  for ATLAS and CMS,  $2\text{--}5 \times 10^{32}$  cm $^{-2}$ s $^{-1}$  for LHCb, and  $3 \times 10^{30}$  cm $^{-2}$ s $^{-1}$  for ALICE). ATLAS and CMS have similar acceptance for beauty measurements. On one hand, their minimum accessible  $p_T$  is larger than for ALICE because of the strong magnetic fields and the larger material budget in the inner tracking detectors; on the other hand, the strong magnetic fields, together with the high luminosity, allow those experiments to cover transverse momenta up to 200–300 GeV/ $c$ . In terms of acceptance for beauty measurements, ALICE overlaps with ATLAS and CMS at central rapidity and with LHCb at forward rapidity. The moderate magnetic field allows measurements down to transverse momenta of less than 1 GeV/ $c$  for charmed and beauty hadrons in the central barrel<sup>1</sup>.

For the performance study presented here, we assume the baseline heavy-flavour production cross sections and yields presented in the ALICE Physics Performance Report, Volume II [13]. Those values are obtained from the pQCD calculations at fixed next-to-leading-order (FO NLO) implemented in the HVQMNR program [16]; note that the cross sections have a theoretical uncertainty of about a factor 2 [13].

### 2.2.1 Exclusive charm meson reconstruction

Among the most promising channels for open charm detection are the  $D^0 \rightarrow K^- \pi^+$  ( $c\tau \approx 120$   $\mu\text{m}$ , branching ratio  $\approx 3.8\%$ ) and  $D^+ \rightarrow K^- \pi^+ \pi^+$  ( $c\tau \approx 300$   $\mu\text{m}$ , branching ratio  $\approx 9.2\%$ ) decays. The detection strategy to cope with the large combinatorial background from the underlying event in Pb–Pb is based on the selection of displaced-vertex topologies [13,

<sup>1</sup>The study of the channel  $B \rightarrow J/\psi + X$ , discussed in section 2.3, should allow a determination of the  $p_T$  differential cross section of B hadrons down to  $p_T \approx 0$ .

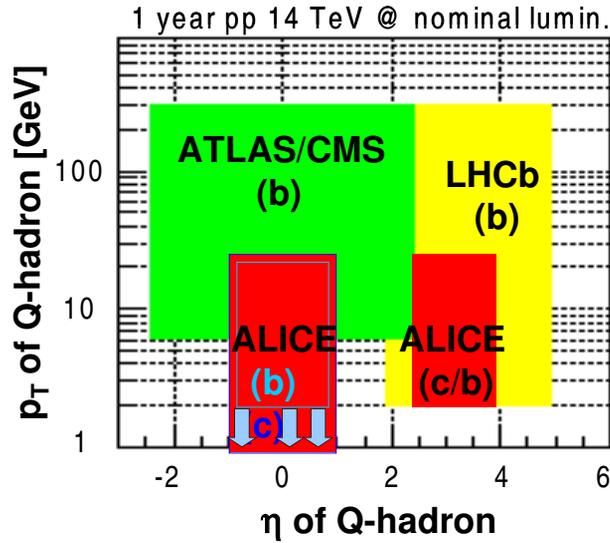


Fig. 4: Schematic acceptance in transverse momentum and pseudorapidity for open heavy flavour hadrons (indicated as Q-hadrons) in the four LHC experiments. The high- $p_T$  coverages correspond to one year (i.e. 7 months) of running at nominal luminosity.

17]. An invariant-mass analysis is used to extract the raw signal yield, to be then corrected for selection and reconstruction efficiency and for detector acceptance. As shown in figure 5 (left), the accessible  $p_T$  range for the  $D^0$  is 1–20 GeV/ $c$  in Pb–Pb and 0.5–20 GeV/ $c$  in proton–proton, with statistical errors better than 15–20% at high  $p_T$ . A similar performance is expected for the  $D^+$  (right-hand panel), though at present the statistical errors are estimated only in the range  $1 < p_T < 8$  GeV/ $c$ . The systematic errors (acceptance and efficiency corrections, centrality selection for Pb–Pb) are estimated to be smaller than 15%.

### 2.2.2 Beauty detection via displaced electrons

Beauty detection via electron-identified tracks with a displacement with respect to the primary collision vertex is favoured by the large semi-electronic branching ratio (b.r.  $\approx 11\%$  [18]) and by the significant mean proper decay length ( $c\tau \approx 500 \mu\text{m}$  [18]) of beauty hadrons. The main sources of background for the signal of beauty-decay electrons are: decays of primary D mesons, which have a branching ratio of  $\approx 10\%$  in the semi-electronic channels [18], and have an expected production yield larger by a factor about 20 with respect to B mesons ( $N^{c\bar{c}} \approx 1.6 \times 10^{-1}/\text{ev}$  and  $N^{b\bar{b}} \approx 7.2 \times 10^{-3}/\text{ev}$  [13]); di-electron decays of vector mesons ( $\rho$ ,  $\omega$ ,  $\phi$ ) and Dalitz decays of pion and  $\eta$  mesons (e.g.,  $\pi^0 \rightarrow \gamma e^+ e^-$ ); conversions of photons in the beam pipe or in the inner layers of the ITS; charged pions misidentified as electrons.

Events were generated using PYTHIA [19]. We evaluated the required statistics at about  $10^7$  proton–proton minimum-bias events at  $\sqrt{s} = 14$  TeV,  $10^6$  proton–proton events containing a

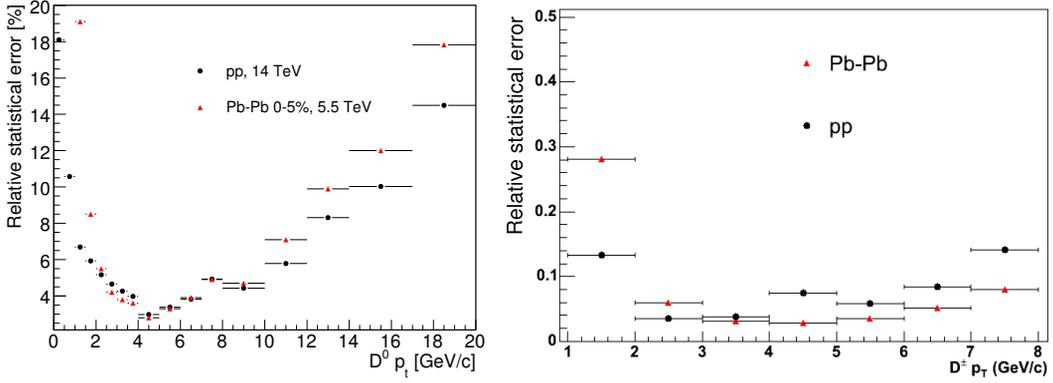


Fig. 5: Expected relative statistical errors for the measurement in ALICE of the production cross sections of  $D^0$  in the  $K^- \pi^+$  channel (left) and  $D^+$  in the  $K^- \pi^+ \pi^+$  channel (right), in 0–5% central Pb–Pb collisions and in proton–proton collisions.

$c\bar{c}$  pair and  $10^6$  proton–proton events containing a  $b\bar{b}$  pair. For the background, we used a sample of  $6 \times 10^6$  minimum-bias proton–proton events. For the proton–proton events with a heavy-quark pair, we used the same PYTHIA settings as for the minimum-bias events, without forcing heavy-flavour production, but selecting events containing a  $c\bar{c}$  or  $b\bar{b}$  pair (in order to obtain a realistic underlying-event multiplicity). Since the resulting shapes of the charm and beauty quarks  $p_T$  distributions are different from those given by NLO pQCD predictions [16] we reweighted the decay electrons in order to match the baseline shapes. The samples, for background and for proton–proton events with a heavy-quark pair, were normalized to one proton–proton event.

Figure 6 shows the distributions of the signal and of the different background sources, in impact parameter, defined in the plane transverse to the beam direction, ( $d_0$ ) and in transverse momentum ( $p_T$ ). The detection strategy is adapted from that developed for Pb–Pb collisions [20] and is based on three steps:

1. Electron identification. Electrons can be efficiently separated from hadrons by combining the PID capabilities of the TPC, and of the TRD. Here we assume for the proton–proton case the same electron PID performance as expected in Pb–Pb collisions [20]. Under the assumption of  $e_{eff}^{TRD} = 90\%$  electron identification probability, the TRD is expected to reject 99% of the charged pions ( $\pi_{eff}^{TRD} = 10^{-2}$  misidentification probability) and fully reject heavier charged hadrons, for  $p > 1$  GeV/ $c$ . Using the information from the TPC, the probability of pion misidentification can be further reduced by a factor of a hundred at low momentum. As the momentum increases and charged pions approach the Fermi plateau in  $\frac{dE}{dx}$ , the additional pion rejection from the TPC decreases and becomes marginal at  $p \simeq 10$  GeV/ $c$ .
2. Primary vertex reconstruction. Due to the TPC and SDD drift speed limitations, during LHC proton–proton runs, the luminosity at the ALICE interaction point has to be kept below  $\mathcal{L}_{max} \simeq 3 \times 10^{30}$  cm $^{-2}$ s $^{-1}$  [12]. When the LHC luminosity will be larger than this value (the design luminosity is about a factor  $10^4$  higher), the luminosity at the ALICE

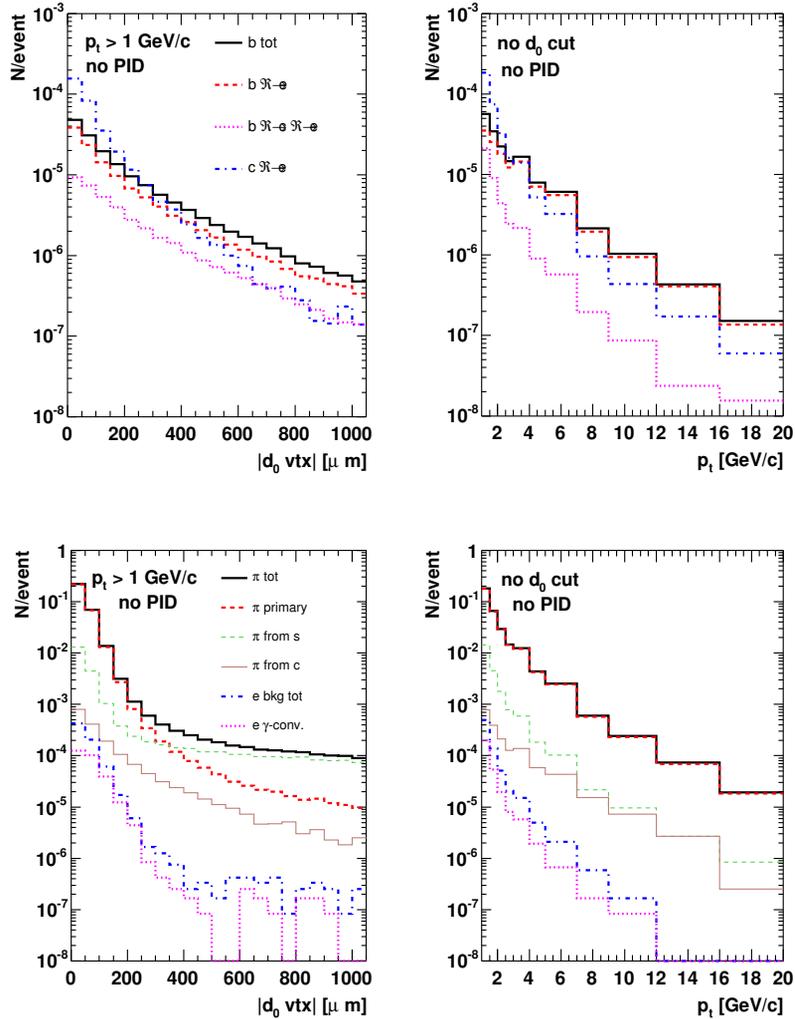


Fig. 6: Beauty and charm decay electrons (top), electrons from other sources and charged pions (bottom), as a function of  $|d_0|$  (left) and  $p_T$  (right). Here,  $|d_0|$  is calculated with respect to the true primary vertex position, known from simulation.

interaction point will have to be reduced, for instance by defocusing the beams, i.e. by enlarging their transverse size up to  $\sigma_{x,y} \sim 150 \mu\text{m}$ . The primary vertex position will be reconstructed on an event-by-event basis, using measured tracks, with an expected resolution of about  $70 \mu\text{m}$  in  $x$  and  $y$  on average [21].

3. Impact parameter cut. Because of the large mean proper decay length ( $\approx 500 \mu\text{m}$ ) of beauty mesons, their decay electrons have a typical impact parameter of a few hundred microns with respect to the primary vertex. A cut  $|d_0| \gtrsim 200 \mu\text{m}$  allows to reject a large fraction of the background (see Fig. 6). We have optimized the value of this cut as a

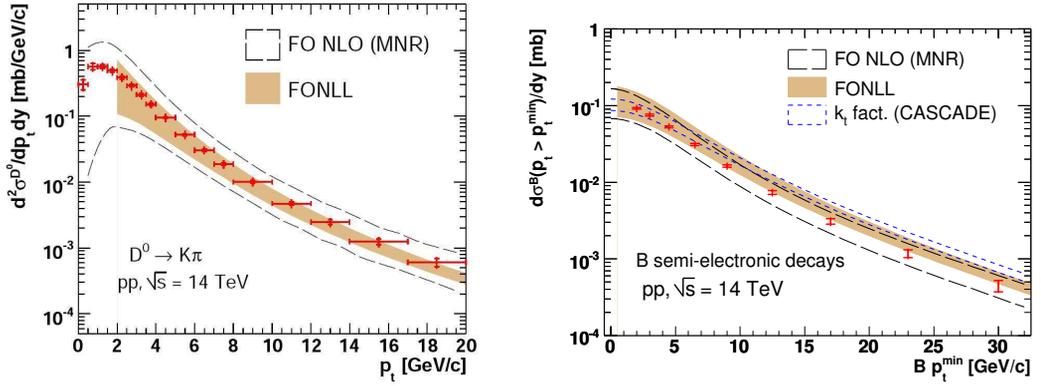


Fig. 7: Differential cross section for  $D^0$  (left) and B meson (right) production as it can be measured with  $10^9$  p–p minimum-bias events. Statistical errors (inner bars) and quadratic sum of statistical and  $p_T$ -dependent systematic errors (outer bars) are shown; the 9% normalization error is not shown. The theoretical predictions from the three pQCD calculations (see text), with their uncertainties are also shown for comparison with the expected experimental sensitivity.

function of the transverse momentum in order to minimize the total errors (statistical and systematic).

To extract the electrons cross section we first apply our cuts to the “measured” electrons. On this selected sample we subtract the residual background (estimated from the charm and pions measurements) and we apply the corrections for efficiency and acceptance. We infer the  $p_T^{\text{min}}$ -differential cross section for beauty mesons,  $d\sigma^B(p_T > p_T^{\text{min}})/dy$ , from the beauty electrons cross section using a procedure similar to that developed by the UA1 Collaboration [22]. The method, described in detail in Refs. [13, 20, 23], is based on Monte Carlo simulation and relies on measured B meson decay kinematics.

### 2.2.3 Results

Figure 7 presents the expected ALICE performance for the measurement of the  $p_T$ -differential cross section of  $D^0$  mesons (left) and the  $p_T^{\text{min}}$ -differential cross section of B mesons,  $d\sigma^B(p_T > p_T^{\text{min}})/dy$  vs.  $p_T^{\text{min}}$  averaged in the range  $|y| < 1$ . For illustration of the sensitivity in the comparison to pQCD calculations, we report in the same figure the predictions and the theoretical uncertainty bands from three approaches [24]: collinearly-factorized FO NLO, as implemented in the HVQMNR code [16], Fixed Order Next-to-Leading Log (FONLL) [25] and  $k_t$ -factorization, as implemented in the CASCADE code [26]. It can be seen that the expected ALICE performance for  $10^9$  events will provide a meaningful comparison with pQCD predictions.

## 2.3 Beauty in the $J/\psi$ channel

Simulation studies are in progress to prepare a measurement of the fraction of  $J/\psi$  that feed-down from B decays. Such measurement can be performed by studying the separation of the dilepton

pairs in the  $J/\psi$  invariant-mass region from the main interaction vertex. The analysis should provide a measurement of the beauty  $p_T$ -differential cross section down to  $p_T \approx 0$ . The pseudo-proper decay time,  $x = L_{xy} \cdot M(J/\psi)/p_T$ , where  $L_{xy}$  is the signed projection of the  $J/\psi$  flight distance on its transverse direction,  $L_{xy} = \vec{L} \cdot \vec{p}_t(J/\psi)/|p_T|$ , can be used to separate  $J/\psi$  from the B decay products from that of prompt decays, as shown in figure 8 for proton–proton collisions. In this expression, the  $M(J/\psi)$  is taken as the known  $J/\psi$  mass [18].

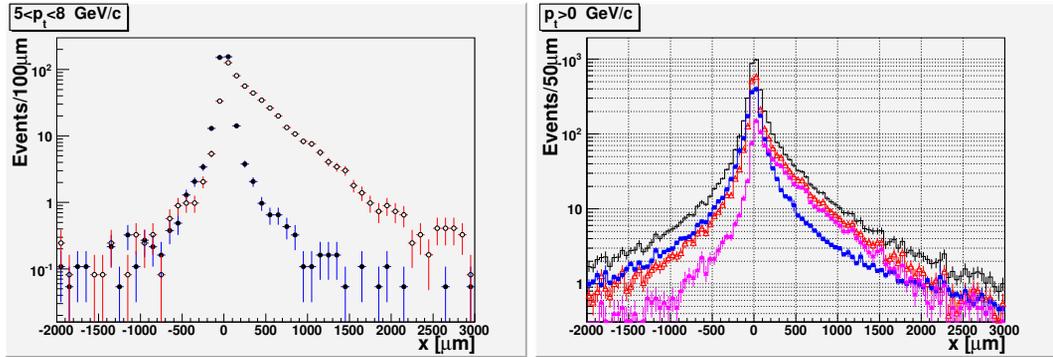


Fig. 8: Distributions of the  $x$  variable, defined in the text, for  $5 < p_T < 8$  GeV/c (left) showed for secondary (open circles) and prompt (closed circles)  $J/\psi$ , and for  $p_T > 0$  (right) showed for total  $J/\psi$  (open triangles), secondary  $J/\psi$  (closed triangles), total background (closed squares) and their sum (lines, in black).

## 2.4 Quarkonia detection in the ALICE Muon Spectrometer

The ALICE experiment will detect heavy quarkonia both at central rapidity in the di-electronic decay channel and at forward rapidity in the di-muonic one. The latter channel will be discussed here.

Quarkonia cross sections at LHC energies are provided by the Color Evaporation Model (CEM) [27]. In this model, the quarkonia cross section is the product of the  $Q\bar{Q}$  cross section times a transition probability ( $F_C$ ) which is specific to each state ( $C$ ) but independent of the energy. Ground state cross sections are the sum of direct production and feed-down from higher mass resonances below the thresholds for heavy-flavoured meson production ( $H = B$  or  $D$ ). At leading order:

$$\sigma_C^{CEM} = F_C \sum_{i,j} \int_{4m_Q^2}^{4m_H^2} d\hat{s} \int dx_1 dx_2 f_{i/A}(x_1, \mu^2) f_{j/B}(x_2, \mu^2) \hat{\sigma}_{ij}(\hat{s}) \delta(\hat{s} - x_1 x_2 s) \quad (1)$$

where  $A$  and  $B$  can be any hadron or nucleus,  $ij = q\bar{q}$  or  $gg$ ,  $\hat{\sigma}_{ij}(\hat{s})$  is the  $ij \rightarrow Q\bar{Q}$  subprocess cross section and  $f_{i/A}(x_1, \mu^2)$  is the parton density in the hadron or nucleus. The predictions for proton–proton collisions at 14 TeV are summarized in Table 1. The transverse momentum distributions are obtained by extrapolating the distributions measured by the CDF experiment at  $\sqrt{s} \sim 2$  TeV [28] at the LHC energies. The  $p_T$  and rapidity distributions of  $J/\psi$  and  $\psi(2S)$  from

	$J/\psi$	$\psi(2S)$	$\Upsilon$	$\Upsilon(2S)$	$\Upsilon(3S)$
$\sigma \times \text{BR} (\mu\text{b})$	3.18	0.057	0.028	0.007	0.0042

Table 1: CEM cross sections for quarkonia production in proton–proton collisions at 14 TeV. Cross sections include feed-down from higher mass resonances and branching ratios in lepton pairs.

the decay of B mesons are generated with PYTHIA [19]. The background, consisting in opposite sign dilepton pairs from the decay of charm, beauty, pions and kaons, is produced with PYTHIA as well.

The dimuon invariant-mass yield expected in one year of data taking in proton–proton collisions at the LHC, with a luminosity of  $3 \times 10^{30} \text{ cm}^{-2}\text{s}^{-1}$  is shown in figure 9. The left (right) panel shows the results obtained with a trigger  $p_T$  cut of 1 (2) GeV/ $c$  in the  $J/\psi$  ( $\Upsilon$ ) mass region. About  $2.8 \times 10^6$   $J/\psi$  and  $2.7 \times 10^4$   $\Upsilon$  are expected. The high statistics allows to reconstruct differential distributions with a fine binning, as illustrated in figure 10.

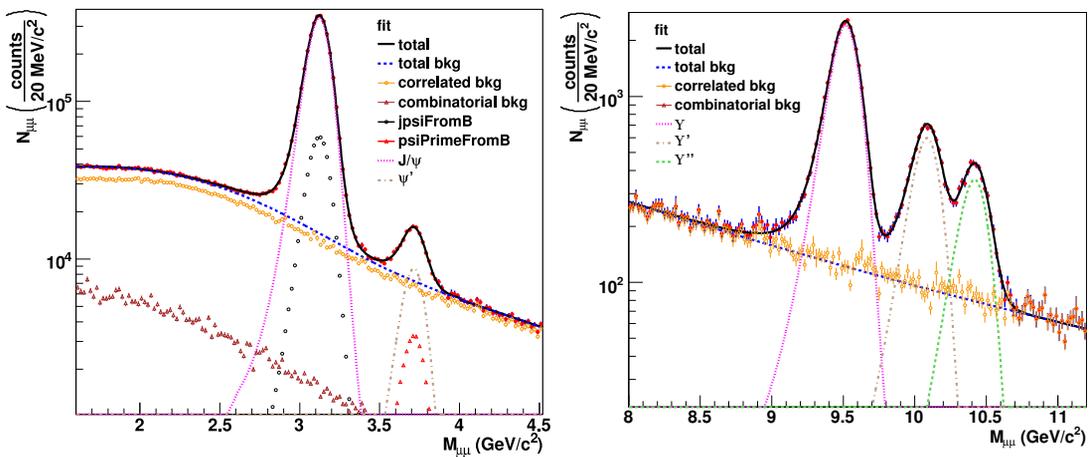


Fig. 9: Dimuon invariant-mass distribution expected in 1 year of data taking in proton–proton collisions at  $3 \times 10^{30} \text{ cm}^{-2}\text{s}^{-1}$ .

#### 2.4.1 Sensitivity to low $x$ PDFs

The choice of a hadron collider implies some uncertainties in the determination of the initial state during the collisions, which are related to the composite nature of the colliding particles. At high energies the hadrons do not interact as a whole: the scatterings takes place between the constituent quarks and gluons. An accurate knowledge of the momentum distribution of such elementary particles in the hadrons is a fundamental issue. The mapped phase space is constantly increased by taking into account data from experiments at different energies. At present, the gluon distribution are constrained by measurements down to  $x$  values higher than  $10^{-4}$ , and extrapolated down to about  $10^{-5}$ .

Leading order calculations show that in proton–proton collisions at 14 TeV, the  $J/\psi$  with a rapidity higher than 3 are produced by gluons with  $x < 10^{-5}$ . Figure 11 shows a comparison

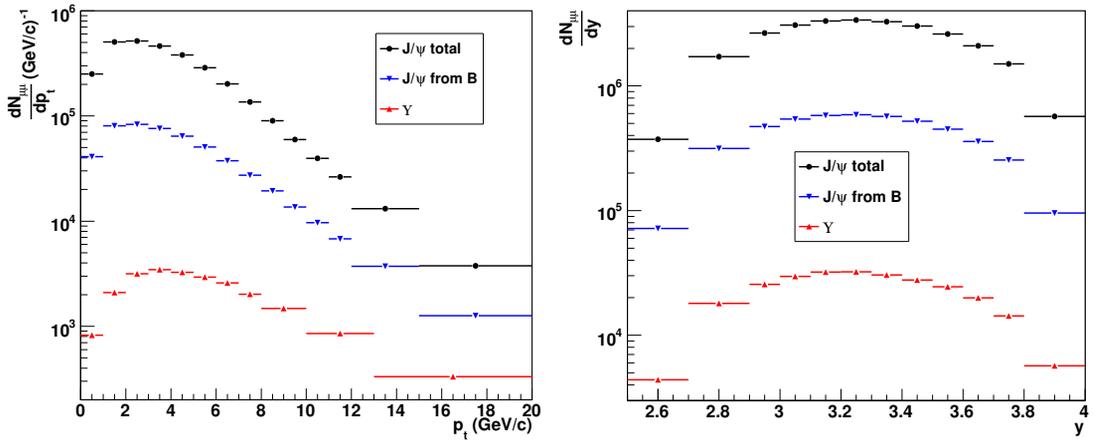


Fig. 10: Expected  $p_T$  (left panels) and rapidity (right panels) differential yields for  $J/\psi$  and  $\Upsilon$  in 1 year of data taking in proton–proton collisions at 14 TeV at the ALICE nominal luminosity of  $3 \times 10^{30} \text{ cm}^{-2} \text{ s}^{-1}$ .

among the PDF sets calculated at Leading Order by the collaborations Martin-Roberts-Stirling-Thorne (MRST98 [29] and MRST01 [30]) and the Coordinated Theoretical-Experimental Project on QCD (CTEQ5 [31] and CTEQ6 [32]) at the scale of  $J/\psi$  (left panel) and  $\Upsilon$  (right panel). For the MRST01 and CTEQ5 sets, two different extrapolations in the low  $x$  region are shown. Differently from the  $\Upsilon$  case, the  $x$ -values explored by  $J/\psi$  in the ALICE Muon Spectrometer acceptance (in yellow), partially sit on the region of extrapolation.

Performing Leading Order calculations in the framework of the Color Evaporation Model,

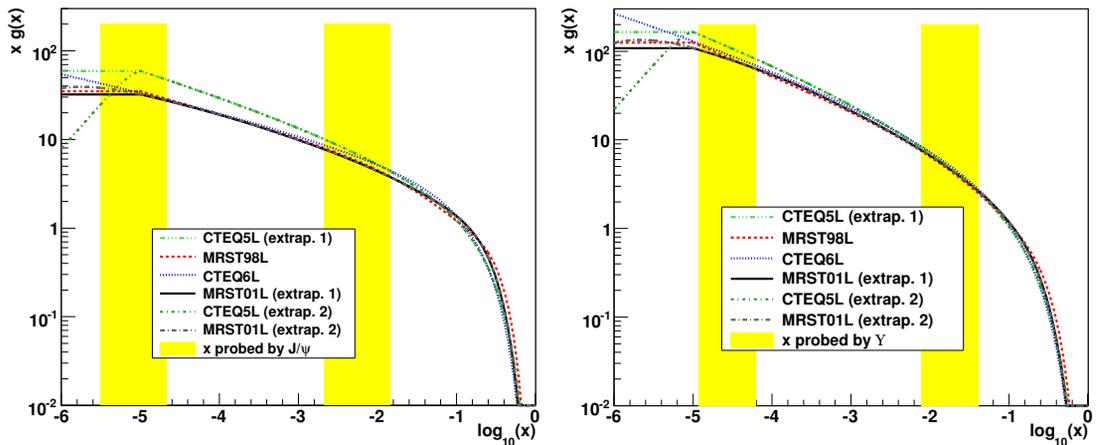


Fig. 11: Comparison of the gluon distributions from MRST and CTEQ.  $x$  regions probed by  $J/\psi$  and  $\Upsilon$  produced in proton–proton collisions at  $\sqrt{s} = 14$  TeV in the rapidity region  $2.5 < y < 4$  are shown.

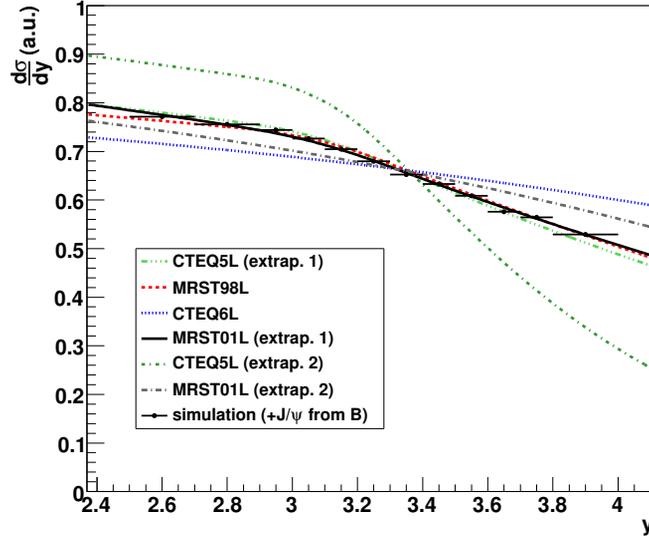


Fig. 12: Comparison between  $J/\psi$  rapidity distributions obtained with four different PDF sets (area in detector acceptance is normalized to 1). Simulation results are also shown.

it is possible to derive the  $J/\psi$  rapidity distribution:

$$\frac{d\sigma_{J/\psi}^{CEM}}{dy} = \frac{F_{J/\psi}}{s} \sum_{i,j} \int_{4m_Q^2}^{4m_H^2} d\hat{s} \hat{\sigma}_{ij}(\hat{s}) f_{i/A}(\sqrt{\frac{\hat{s}}{s}}e^y, \mu^2) f_{j/B}(\sqrt{\frac{\hat{s}}{s}}e^{-y}, \mu^2) \quad (2)$$

Figure 12 shows a comparison of the differential distributions obtained with different sets of PDFs. It is worth noting that such distributions are normalized by setting equal to unit their integral from 2.5 to 4 rapidity units. The *shape* of the distributions is clearly dependent on the behavior of the gluon functions: results obtained with MRST98L, MRST01L (extrap. 1) and CTEQ5L (extrap. 1), which were extrapolated flat in  $xg(x)$  (figure 11) are compatible among each other and are clearly different from the result obtained with CTEQ6L, and with a changed extrapolation behavior of MRST01L (extrap. 2) and CTEQ5L (extrap. 2). The results show that also a small change in the PDFs extrapolation (see MRST01L) can lead to appreciable changes in the shape of the differential distributions.

The comparison between the simulation results and the calculations (figure 12), show that, due to the high statistics, the accuracy of the data expected to be collected by the ALICE Muon Spectrometer will be good enough to allow to discriminate among different shapes of the gluon distribution functions in the region of  $x < 10^{-5}$  (at least in the frame of a leading order analysis).

## 2.5 Conclusions

We presented the expected performance of ALICE for the study of open heavy flavour and quarkonium states in nucleus–nucleus collisions at the LHC. Thanks to the good expected performance of the detectors for tracking and particle identification, and its low magnetic field in

the central barrel, the ALICE acceptance is complementary to that of other LHC experiments. This opens the possibility of measuring heavy-flavour production down to  $p_T \approx 1 \text{ GeV}/c$ , both at central and forward rapidity. Our results indicate that ALICE can provide several  $p_T$  differential measurements of charmed and beauty hadrons production with errors that are smaller or comparable to the theoretical uncertainties of pQCD calculations. The invariant-mass resolution of the muon spectrometer allows to resolve several charmonium and bottomonium resonances and the statistics expected to be collected in the first year of data taking should be enough to study in details the production of several quarkonia states. Finally, ALICE will probe the parton distribution functions down to unprecedentedly low values of the Feynman  $x$  variable.

### 3 Quarkonia and open beauty production in the ATLAS experiment at the LHC

*Authors: E. Lytken and M. zur Nedden*

#### 3.1 Introduction

ATLAS [33] is a general-purpose experiment with main emphasis on searches for new phenomena based on high  $p_T$  particles. Since most of the  $B$ -physics appears in the lower  $p_T$  range, triggering within the LHC environment on those events is a challenge. Nevertheless, ATLAS has good capabilities for a rich  $B$ -physics program, based on the dedicated and flexible trigger, the precise and flexible vertexing and tracking, the good muon identification and the high-resolution calorimetry. Furthermore, theoretical descriptions of heavy flavoured hadrons need input from the LHC, where precision measurements are already achievable after one year of data taking. The expected inclusive production cross-section for  $b\bar{b}$  pairs at LHC is estimated to be  $\sigma_{b\bar{b}} \approx 500 \mu\text{b}$  leading to more than  $10^6$  produced pairs per second at design luminosity. The experimental precision reached at ATLAS should at least allow the verification of the Standard Model (SM) prediction. In the case of the rare  $B$ -decays, clearly higher luminosity is needed to achieve sensitive upper limits for the indirect beyond the Standard Model (BSM) searches. Therefore, the most relevant part of the ATLAS  $B$ -physics program will take place in the initial phase at lower luminosities with an extension into the high luminosity phase. The envisaged measurements are extending the discovery potential for physics beyond the SM by the measurement of  $CP$  violation parameters, predicted to be small in the SM, and of rare  $B$  decays.

The exclusive  $B^+$  channel provides a clean reference signal. Due to the clear event topology and its rather large branching ratio, it can be measured during the initial luminosity phase of the LHC. The  $B^+ \rightarrow J/\psi K^+$  decay can serve as a reference channel for the measurement of the decay probability of a very rare decay channel  $B_s \rightarrow \mu^+ \mu^-$ , which is strongly suppressed in the Standard Model and therefore offers a good sensitivity to new physics. The total and differential cross-sections of the rare  $B$  decays will be measured relative to the  $B^+ \rightarrow J/\psi K^+$  cross-section allowing thus the cancellation of common systematic errors. Furthermore, it can also act as a control channel for the  $CP$  violation measurement and can be used to estimate the systematic uncertainties and efficiencies of flavour tagging algorithms. Finally, the relatively large statistics for this decay allows for initial detector performance studies.

The trigger menu for the ATLAS  $B$ -physics program has been designed to take maximum advantage of the early run phases at lower luminosities ( $\mathcal{L} < 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ ). Since only 5 - 10 %

of the limited bandwidth of the ATLAS trigger system is devoted to the  $B$ -physics triggers, highly efficient and selective triggers are needed. Most  $B$ -physics triggers are based on single- and di-muon events in the final state leading to a clean signature for triggers and flavour tagging [34–36]. In the early data taking period the main  $B$ -physics triggers are expected to run without a need of prescales, allowing for low  $p_T$  muon and low  $E_T$  electron triggers (the latter will be however prescaled at  $10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ ). In general, the trigger strategy is mainly based on a single muon trigger at the first level, which could be combined with certain calorimeter trigger objects at higher trigger levels to select hadronic final states ( $B_s \rightarrow D_s \pi$ ) or  $e/\gamma$  final states ( $J/\psi \rightarrow e^+ e^-$ ,  $K^* \gamma$  or  $\phi \gamma$ ). In order to not exceed the available bandwidth, in the phase of higher luminosities above  $2 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$  the main working trigger will be based on di-muons on the first level, enabling a clean measurement of rare  $B$ -decays ( $B \rightarrow \mu\mu$  or  $B \rightarrow K^{*0} \mu\mu$ ), double semi-leptonic decays and the  $B \rightarrow J/\psi(\mu\mu)$  decay channels.

### 3.2 Beauty production cross section determination

The  $b\bar{b}$  production cross-section will be measured using inclusive and exclusive methods in parallel to control the systematics. For the inclusive methods ATLAS looks at the semi-leptonic  $b \rightarrow \mu + X$  and the  $B \rightarrow J/\psi(\rightarrow \mu^+ \mu^-) + X$  decay modes. In the next section we will briefly describe the measurement of the exclusive  $B^\pm \rightarrow J/\psi K^\pm$  cross-section. The measurement of the  $J/\psi$  mass and its detection efficiency is a central task for the analysis of the first ATLAS data, providing the tools to validate the detector by extracting muon energy scale determination in the low  $p_T$  region and detector misalignments (Sect. 3.5). Finally, the mass measurement and reconstruction efficiency for  $B^+$ , the total and differential cross-sections and its lifetime measurements will be of interest for other  $B$ -physics analyses.

The main backgrounds that are competing with the signal are single-muon from  $c\bar{c}$  decays and direct  $J/\psi$ 's from  $pp \rightarrow J/\psi + X$ . In the first case, the  $p_T$  distribution of the muons is softer as compared to the muon spectrum from  $b\bar{b}$  decays while in the latter no displaced secondary vertex is expected. In consequence the following parameters are used for  $b$ -tagging:

- the signed transverse impact parameters  $d_0$  of charged particles originating from  $B$ -meson decays at a secondary vertex due to the long lifetime of  $B$ -mesons.
- the relative transverse momentum  $p_T^{\text{rel}}$  of the muon of the  $b$ -decay with respect to the axis of the associated jet.

The measurement of the  $p_T^{\text{rel}}$  distribution of the selected muons offers a good possibility to determine the  $b$ -contents fraction in the offline analysis. In the rest frame of the decaying  $B$ -meson, the muon gets a high transverse momentum, which is significantly larger than in the case of charm or light quark decays. The relative transverse muon momentum,  $p_T^{\text{rel}}$  can therefore be used to determine the  $b$ -content of a selected data sample by fitting Monte Carlo templates to data. For the  $B$ -mesons at ATLAS, generally decay lengths of the order of several mm are expected which are at the same order of magnitude as the expectation for  $D$ -mesons. The signed transverse impact parameter  $d_0$  is a boost independent quantity. For large transverse momenta ( $p_T^\mu > 10 \text{ GeV}$ )  $d_0$  is proportional to the lifetime of the decaying particle and positive values of  $d_0$  are preferable. The significance of signed impact parameter for muons with an associated  $b$ -jet and the distributions of the relative transverse momentum are shown in Fig. 13 and Fig. 14 respectively. In both cases, the selection power is clearly visible.

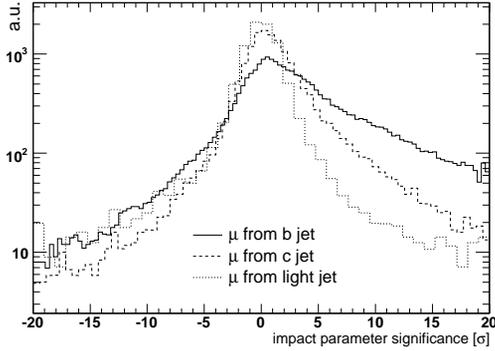


Fig. 13: The distribution of the significance of the signed impact parameter  $d_0/\sigma$

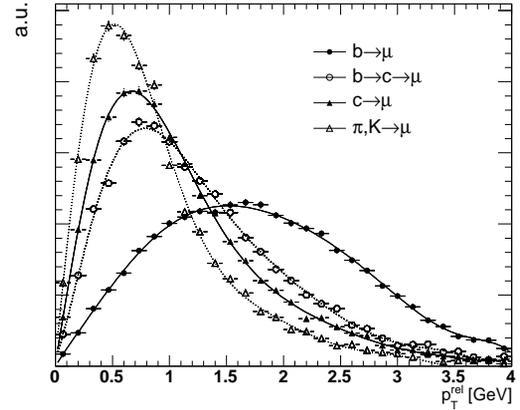


Fig. 14: The  $p_T^{\text{rel}}$  distribution for different processes considered in the  $b$ -jet selection.

The  $b\bar{b}$  production cross-section measurement based on the single-muon and jet requirements at the trigger level is then obtained according to the usual relation

$$\sigma(b\bar{b} \rightarrow \mu(p_T^\mu > 6 \text{ GeV})X) = \frac{N_b^{\text{sel}}}{\int \mathcal{L} dt} \cdot \frac{f_b}{\epsilon_b^{\text{trig}} \cdot \epsilon_b^{\text{rec}}} \quad (3)$$

where the  $b$ -trigger efficiency was found to be  $\epsilon_b^{\text{trig}} = 0.135$  and the combined muon reconstruction efficiency is  $\epsilon_b^{\text{rec}} = 0.85$ . The determination of the  $b$  content,  $f_b$ , of the selected sample is extracted by fitting the simulated  $p_T^{\text{rel}}$  distribution to the data. This can be done by a binned maximum likelihood fit taking into account the finite size of both, the data sample and the simulated Monte Carlo templates. In the signal template, the direct  $b \rightarrow \mu$  and cascade  $b \rightarrow c \rightarrow \mu$  contributions are contained, whereas all others are summarized in the background template. The distribution can be seen in Fig. 15. With this fit, a  $b$ -content of  $f_b = 0.23$  was obtained, and a corresponding background fraction of  $f_{bg} = 0.77$ . The values obtained in this study are in good agreement with the values obtained by the Tevatron experiments [37]. Combining both methods, the  $b\bar{b}$  production cross-section is expected to be measured with a statistical precision better than  $\mathcal{O}(1\%)$  with  $\approx 100 \text{ pb}^{-1}$  of integrated luminosity. The systematic uncertainty is dominated by the luminosity measurement. It is estimated to be 10% in the initial phase, and reduced to about 6.5% after the first  $0.3 \text{ fb}^{-1}$ . The scale uncertainty of the NLO calculations is about 5%, while the PDF uncertainty is estimated to be 3%. Finally, the uncertainty originating from the muon identification is about 3%, leading to a systematic uncertainty of 12% and 9.2% correspondingly in the initial and later phase.

### 3.3 $B^+$ reference channel

Negligible direct  $CP$  violation is expected in the  $B^\pm \rightarrow J/\psi K^\pm$  decay because for  $b \rightarrow c + \bar{c}s$  transitions the SM predicts that the leading and higher order diagrams are characterized by the same weak phase. The only source of asymmetry is the different interaction probability for

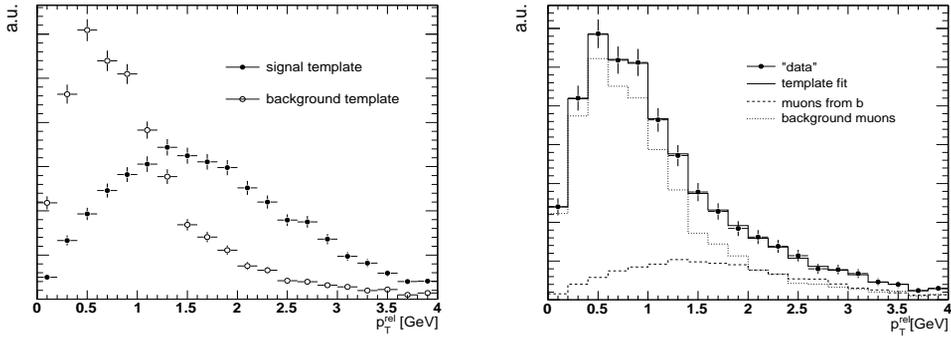


Fig. 15: The generated  $p_T^{\text{rel}}$  templates (left) and fraction of the  $b$ -content of the selected  $b$ -jet sample (right) showing comparison of the fitted to the true values from Monte Carlo.

$K^+$  and  $K^-$  with the detector material. The  $B^+$  candidates are reconstructed based on the  $J/\psi(\rightarrow \mu^+\mu^-)$  selection, and combined with  $K^+$  candidates formed from inner detector tracks.

The  $B^+$  invariant mass distribution  $M(K^+\mu^+\mu^-)$  of the candidates, fulfilling the selection cuts, is presented in Fig. 16 for signal and background with a maximum-likelihood fit, where the likelihood function is a Gaussian for the signal region and a linear function for the background ( $b\bar{b} \rightarrow J/\psi + X$ ). The mass range of the fit is taken from 5.15 GeV to 5.8 GeV in order to reduce contributions from partially reconstructed  $B$  meson decays. The background at the right of the mass peak originates from misidentified  $\pi^+$  from  $B^+ \rightarrow J/\psi\pi^+$  decays. The fit result for the  $B^+$  mass is:  $M(B^+) = (5279.3 \pm 1.1)$  MeV with a width of  $\sigma(B^+) = (42.2 \pm 1.3)$  MeV. The relative errors, scaled properly for an integrated luminosity of about  $10 \text{ pb}^{-1}$ , are about 0.02% and 3.5% respectively. The slight shoulder to the left of the mass distribution is due to the background shape in this mass region and has been included in the systematic uncertainties of the fit model.

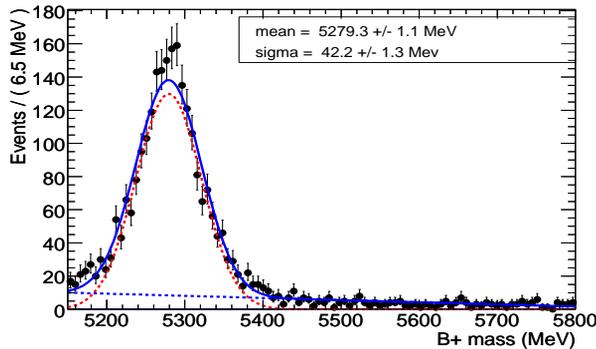


Fig. 16:  $B^+$  mass fit with the both signal (red) and background (blue) contributions shown separately.

With the first  $10 \text{ pb}^{-1}$  of LHC data a total and differential production cross-section measurement of the  $B^+ \rightarrow J/\psi K^+$  can be achieved. The differential cross-section  $d\sigma/dp_T$  can be obtained from the usual form:

$$\frac{d\sigma(B^+)}{dp_T} = \frac{N_{\text{sig}}}{\Delta p_T \cdot \mathcal{L} \cdot \mathcal{A} \cdot \text{BR}} \quad (4)$$

where  $N_{\text{sig}}$  is the number of reconstructed  $B^+$  obtained from the mass fit and the size of the  $p_T$  bin is denoted with  $\Delta p_T$ . Furthermore,  $\mathcal{L}$  is the total luminosity to which the dataset corresponds and is obtained from PYTHIA output and BR is the product of the branching ratios using the world average [38] branching ratios of  $\text{BR}(B^+ \rightarrow J/\psi K^+) = (10.0 \pm 1.0) \times 10^{-4}$  and  $\text{BR}(J/\psi \rightarrow \mu^+ \mu^-) = (5.88 \pm 0.10) \times 10^{-2}$ . The overall efficiency  $\mathcal{A}$  is calculated for each  $p_T$  range separately as the ratio of the number of signal events determined from the previous fit procedure and the number of the Monte Carlo signal events within the same  $p_T$  range. To measure the  $B^+$  total cross-section a similar procedure with that used for the calculation of the differential cross-section is followed.

The measurement of the lifetime  $\tau$  of the selected  $B^+$  candidates is a sensitive tool to confirm the beauty content in a sample, in particular the number of the reconstructed  $B^+ \rightarrow J/\psi K^+$  decays obtained in the  $b\bar{b} \rightarrow J/\psi X$  dataset. The proper decay-time is defined as  $t = \lambda/c$ . The proper decay-time distribution in the signal region  $B^+ \rightarrow J/\psi K^+$  can be parametrized as a convolution of an exponential function with a Gaussian resolution function, while the background distribution parametrization consists of two different exponential functions, where each is convoluted with a Gaussian resolution function. In the  $b\bar{b} \rightarrow J/\psi X$  no zero lifetime events are expected since there is no prompt  $J/\psi$  produced. In the realistic case, where zero lifetime events will be present, an extra Gaussian centered at zero is needed in order to properly describe those events.

The results on the lifetime measurements are shown in Fig. 17. The background can be best described with the two lifetime components ( $\tau_1$  and  $\tau_2$ ). For the events in the mass region of the signal within  $M(B^+) \in [5.15, 5.8] \text{ GeV}$  the proper decay-time found from the decay length is compared to the generated  $B^+$  lifetime. The differences are well centered at zero with a Gaussian distribution and sigma  $0.088 \text{ ps}$ . It should be noted that the resolution as well as its  $\sigma$  in  $\eta$  bins of  $0.25$  is found to be independent of  $\eta$ .

### 3.4 Open flavour: rare B-decays

Flavour changing neutral currents, a direct transition from  $b \rightarrow d/s$ , are forbidden at the tree level in the SM and occur at the lowest order through one loop diagrams. They are a sensitive test of the SM and its possible extension(s), providing information on the long distance QCD effects and enabling a determination of the CKM matrix elements  $|V_{td}|$  and  $|V_{ts}|$ . Furthermore, some of the rare decay channels contribute to the background for other channels, which are very sensitive to BSM effects.

An upper limit of the branching ratio  $BR(B_s^0 \rightarrow \mu^+ \mu^-) = (1 - 2) \cdot 10^{-8}$  at 90 % confidence level or of  $(2 - 3) \cdot 10^{-8}$  at  $3\sigma$  evidence based on  $N_B = 1.1$  events that can already be extracted from an integrated luminosity of  $2 \text{ fb}^{-1}$ . This is clearly better than the current CDF limit of  $4.7 \cdot 10^{-8}$  at 90 % confidence level. Already at  $1 \text{ fb}^{-1}$  ATLAS is able to collect  $\mathcal{O}(10^6)$

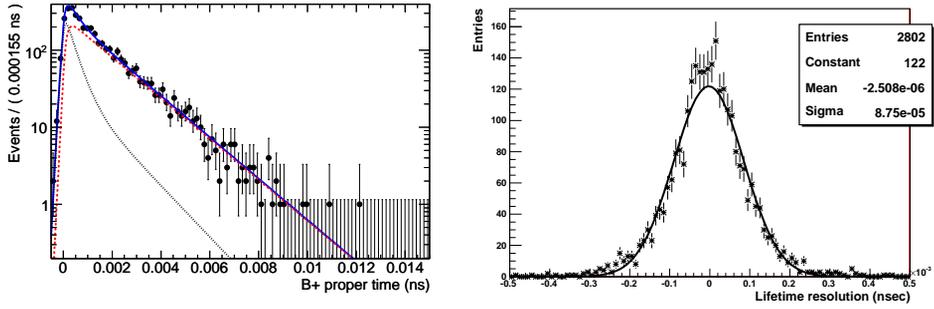


Fig. 17: The  $B^+$  lifetime fit (left) with the signal (dashed red) and the background (dashed black) contributions shown separately. The  $B^+$  lifetime resolution (right).

di-muon events in the mass window of  $4 \text{ GeV} < M(\mu^+\mu^-) < 7 \text{ GeV}$ . This is after the selection based on cuts on  $p_T$ , the invariant mass  $M_{\mu^+\mu^-}$ , the transverse decay length  $L_{xy}$  of the di-muon system, and on isolation requirements. Based on this data, an upper limit on the number of signal events,  $N_B$ , corresponding to a given confidence level will be determined. The main background sources originate from combinatorial decays as  $b\bar{b} \rightarrow \mu^+\mu^-X$ , from misidentifications ( $B_s^0 \rightarrow \pi^+\pi^-$ ,  $B_s^0 \rightarrow K^+K^-$ ,  $B_s^0 \rightarrow \pi^+K^-\nu_\mu$ ) or from other rare decays ( $B_s^0 \rightarrow \mu^+\mu^-\mu^+\nu_\mu$ ,  $B_s^0 \rightarrow \mu^+\mu^-\gamma$ ).  $N_B$  will be used to extract the upper limit on the  $B_s^0 \rightarrow \mu^+\mu^-$  branching ratio  $BR(B_s^0 \rightarrow \mu^+\mu^-)$ , using the reference channel  $B^+ \rightarrow J/\psi K^+$  as described in Sec. 3.3, since trigger and offline reconstruction efficiencies largely cancel for di-muons in these channels. In this procedure, a ratio of geometric and kinematical acceptances of the signal and the reference channel will be determined from the Monte Carlo simulations. With an integrated luminosity of  $30 \text{ fb}^{-1}$ , corresponding to three years of initial data taking, the SM predictions could be tested with a  $3\sigma$  sensitivity. The continuation of this measurement at nominal LHC luminosities has been proved to lead to a clear statement with a  $5\sigma$  sensitivity after already one additional year of data taking at design luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ .

### 3.5 Quarkonia

Understanding the production of prompt quarkonia at the LHC is an important step to understand the underlying QCD mechanisms, and one that has given rise to controversy, both with respect to the cross-section magnitude [39] and the polarization [40]. The initial discrepancy in cross-section led to the Color Octet Model [41] but more high  $p_T$  results are needed to distinguish between this and competing models.

In addition to these open questions, the narrow  $J/\psi$  and  $\Upsilon$  resonances are ideal for studies of detector performance. The expected abundant production (see Fig. 18) makes this feasible already in the very early data. Both decay channels  $J/\psi$  ( $\Upsilon$ )  $\rightarrow \mu^+\mu^-$  and  $J/\psi$  ( $\Upsilon$ )  $\rightarrow e^+e^-$  will be used as tools to test our detector performance. In the following we consider only the  $J/\psi$  and  $\Upsilon(1S)$  resonances. Quarkonia selection in ATLAS is mainly based on a di-muon trigger which requires two identified muons, both with  $p_T \geq 4 \text{ GeV}$  and within a pseudorapidity of

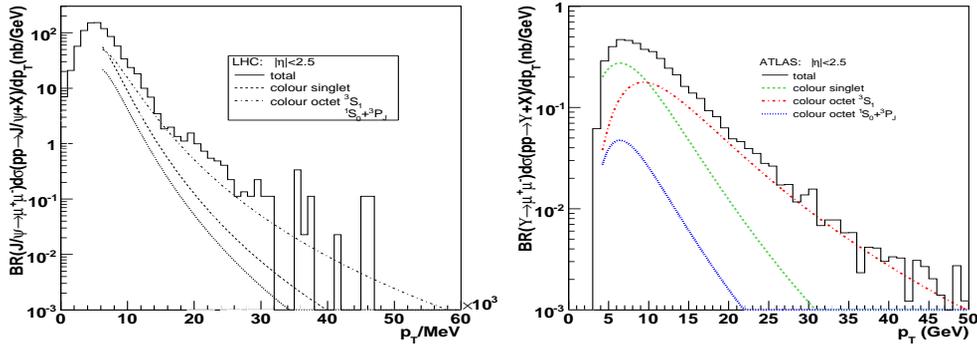


Fig. 18: Differential  $J/\psi$  and  $\Upsilon$  cross-sections as predicted from the Color Octet Model. Contributions from (singlet)  $\chi$  production is included.

$|\eta| < 2.4$ . The di-muon sample considered here has offline  $p_T$  cuts of 6 and 4 GeV applied to the two identified muons. To suppress backgrounds (decays-in-flight, heavy flavour decays) we require tracks to come from the same vertex and with a pseudo-proper time cut of  $\tau = 0.2$  ps, defined as  $\tau = \frac{M \cdot L_{xy}}{p_T(J/\psi) \cdot c}$ . In Fig. 19 (left) the resulting di-muon spectrum with background contributions is shown. We expect 15000  $J/\psi$ 's and 2500  $\Upsilon(1S)$  per  $\text{pb}^{-1}$ . The mass resolution for  $J/\psi \rightarrow \mu^+\mu^-$  is expected to be 53 MeV, and for  $\Upsilon \rightarrow \mu^+\mu^-$  we found 161 MeV on average.

We are also studying the possibility of doing performance measurements using di-electron resonances. In that case the  $E_T$  cut for both leptons is 5 GeV at trigger level and offline, and  $|\eta| < 2$ . Tight electron identification cuts are applied to reject background, including  $E/p$ , vertexing layer hit on the tracks, and the ratio of high to low threshold hits in the transition radiation tracker. We expect 2500  $J/\psi$ 's and 500  $\Upsilon \rightarrow e^+e^-$  per  $\text{pb}^{-1}$  with an instantaneous luminosity of  $10^{31} \text{ cm}^{-2} \text{ s}^{-1}$ . The mass resolution for  $J/\psi \rightarrow e^+e^-$  is expected to be about 200 MeV, see Fig. 19 right. The width is mainly constrained by bremsstrahlung due to the large amount of material in the inner detector.

In addition to cross-section measurements ATLAS will use the quarkonia di-muon decays to provide answers to the polarization puzzle and help constrain the models. Defining the polarization parameter  $\alpha$  as  $\alpha = (\sigma_T - 2\sigma_L)/(\sigma_T + 2\sigma_L)$ , we can measure this by  $\theta^*$ , the angle between  $J/\psi$  in rest frame and  $\mu^+$ , as they are related by:

$$\frac{dN}{d \cos \theta^*} = C \cdot \frac{3}{2\alpha + 6} \cdot (1 + \alpha \cos^2 \theta^*) \quad (5)$$

With the di-muon triggers we get a rather narrow  $\cos \theta^*$  distribution, with both muons having similar  $p_T$ . To access higher values of  $\cos \theta^*$  we utilize a single muon trigger where we can pair the trigger muon with a low  $p_T$  track to get large  $\Delta p_T$  and  $\cos \theta^*$  (see Fig. 20 left). For the result quoted here we used a trigger threshold of 10 GeV for the single muon trigger and the  $p_T$  requirement on the second track was 0.5 GeV. The looser cuts allow for more background but still with decent signal to background discrimination ( $S/B = 1.2$  for  $J/\psi$ ). This dataset was added (with corrections for overlaps) to complement the di-muon triggered dataset. The

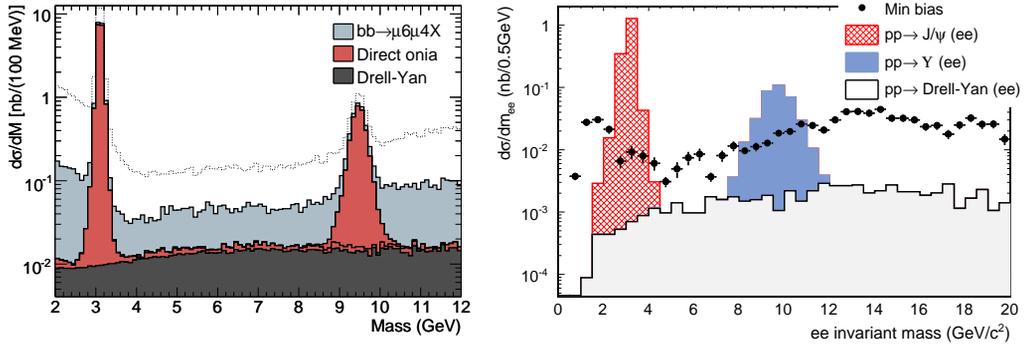


Fig. 19: Di-electron mass distributions for muons (left) and electrons (right). In the left plot the spectrum in case of no vertex cuts (top dashed line) is also shown.

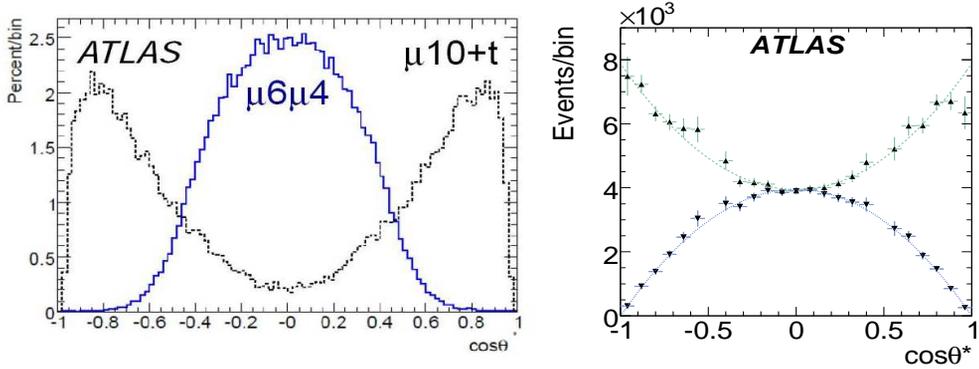


Fig. 20: Left:  $\cos\theta^*$  distributions for double and single muon triggered events. Right: Fit to combined events in  $p_T \in [12, 13]$  GeV. Top shows longitudinally polarized ( $\alpha = -1$ ) and bottom transverse polarized ( $\alpha = 1$ ) events.

combined  $\cos\theta^*$  distributions were then fitted for  $\alpha$  and  $C$  in slices of  $p_T$ . With unpolarized samples the results are given in Table 2 for  $10 \text{ pb}^{-1}$ . Similar tests have been carried out with  $\alpha = \pm 1$ . An example is shown in Fig. 20 right.

We expect to be able to measure the  $J/\psi$  polarization with the  $p_T$  of the  $J/\psi$  in the range of 10 GeV (trigger dependent) up to 50 GeV. Already with the first  $10 \text{ pb}^{-1}$  we can achieve better precision than the current Tevatron measurements - but with  $J/\psi$  at high  $p_T$ , which is what is needed to truly distinguish between models. For  $\Upsilon(1S)$  we can get the same precision with  $100 \text{ pb}^{-1}$ . In this latter case we have acceptance all the way down to  $p_T \approx 0$  for the  $\Upsilon$ , which will be a very useful region to compare with the Tevatron results.

Table 2:  $10 \text{ pb}^{-1}$ : Measured values of  $\alpha$  in  $p_T$  bins in simulated datasets with  $\alpha = 0$ .

$p_T$ (GeV/c)	9 - 12	12 - 13	13 - 15	15 - 17	17 - 21	> 21
$\alpha(J/\psi)$	0.156 $\pm 0.166$	-0.006 $\pm 0.032$	0.004 $\pm 0.029$	-0.003 $\pm 0.037$	-0.039 $\pm 0.038$	0.019 $\pm 0.057$
$\alpha(\Upsilon)$	-0.42 $\pm 0.17$	-0.38 $\pm 0.22$	-0.200 $\pm 0.20$	0.08 $\pm 0.33$	-0.15 $\pm 0.18$	0.47 $\pm 0.22$

### 3.6 Conclusions

The ATLAS experiment has a rich  $B$ -physics program [42] based on clearly defined trigger strategies for all luminosity phases of the LHC. These measurements will contribute to  $CP$  violation studies with  $B_s$ -mesons and its sensitivity to BSM as well in studies of rare  $B$  decays and quarkonia production. The precision measurement of  $B$ -physics processes are an alternative method to explore the presence of new physics at LHC in addition to the direct SUSY searches.

Two inclusive methods for beauty cross-section measurements to be used mainly at the early data taking period of ATLAS were presented. The first method is using the  $J/\psi$  signature with detached vertices, while the second one is based on semileptonic  $b \rightarrow \mu$  decays. The two methods are complementary and the plan is to apply them simultaneously since both signatures will be available with early data. The two methods rely on different trigger algorithms and different physics processes and signatures, therefore the cross-section results obtained from each one can be used for cross checking calibrations of the trigger algorithms used in the measurements. Combining these two methods, the  $b\bar{b}$  production cross-sections measurement is expected to reach a statistical precision of  $\mathcal{O}(1 \%)$  after one month of data taking, if the initial LHC luminosity will be  $\mathcal{L} = 10^{31} \text{ cm}^2\text{s}^{-1}$ , whereas a systematic uncertainty of  $\mathcal{O}(12 \%)$  is expected. Furthermore, the reference channel  $B^+ \rightarrow J/\psi K^+$  has been studied and it could be shown, that a lifetime measurement is a good tool to confirm the  $b$  content of the selected sample.

In the first data taking period ATLAS will also measure the  $J/\psi$  and  $\Upsilon$  cross-sections, taking advantage of the favorable trigger situation in the early data. A method to determine the level of polarization is also presented. We expect to measure the  $J/\psi$  polarization to within 0.02 - 0.06 in the first  $10 \text{ pb}^{-1}$ , dependent on the polarization itself.

## 4 Heavy flavour production in the CMS experiment at the LHC

*Authors: M. Biasini and A. Starodumov*

### 4.1 Introduction

There are several reasons why a general purpose detector like CMS designed for high  $p_T$  physics could be efficiently used to study heavy flavour physics. First of all, there will be about  $10^{11}$   $b\bar{b}$  pairs produced at the initial luminosity year of  $10^{32} \text{ cm}^{-2}\text{s}^{-1}$  thanks to the high  $b\bar{b}$  x-section which is  $\sigma_{b\bar{b}} \sim 500 \mu\text{b}$  at  $\sqrt{s} = 14 \text{ TeV}$ . So, very rare decays like  $B_s^0 \rightarrow \mu^+\mu^-$  and  $B_s^0 \rightarrow \gamma\mu^+\mu^-$  could be searched even with the Standard Model (SM) decay rate. In different scenarios

of New Physics (NP) the branching fraction of these decays could be enhanced by orders of magnitude, which makes the observation of these channels even more probable.

From the detector point of view heavy flavour physics is also an attractive field. Thanks to the low  $p_T$  (di)-muon triggers, precise vertex detector and efficient tracker system, the CMS detector is capable of efficiently recognizing and reconstructing specific topologies of  $b$ -decays. On the other hand, the study of  $b$ -jets provides with an important knowledge which might be crucial in searches for Higgs boson and supersymmetric particles. Also, one should not forget that  $b$ - and heavy onia-decay channels provide an excellent calibration opportunity for the vertex and tracker systems.

And finally, at the low luminosity phase there are no expectations to observe Higgs decays, but as it is mentioned above there are plenty of  $b\bar{b}$ . So, first real physics results at LHC could be obtained in the heavy flavour sector. Even at a high luminosity, thanks to efficient (di)-muon trigger, CMS is able to continue such studies.

## 4.2 The CMS detector

The complete description of the CMS detector can be found elsewhere [43]. Here only the main characteristics are mentioned. The CMS detector is a standard general purpose hadron collider detector composed of the following subsystems: vertex and tracker systems, electro-magnetic and hadron calorimeters and muon system. The full length is 22 m and outer diameter is 15 m. The total weight of the detector is 12.5 kton. All subsystems but the muon detector are placed inside a superconducting magnetic coil which is able to reach a 3.8 T-field. In the following, the most crucial subsystems for heavy flavour physics are briefly discussed.

### 4.2.1 Muon system

The CMS muon system is composed of three types of gaseous particle detectors for muon identification. Drift tubes (DT) chambers in a central barrel region and cathode strip chambers (CSC) in two end-cap regions, thanks to a high spatial resolution, are used for position and momentum measurements. Because of their fast response time, both systems are also provide the Level-1 trigger with good efficiency and high background rejection. Resistive plate chambers (RPC), which are placed in both the barrel and end-cap regions, combine an adequate spatial resolution with an excellent ( $\leq 5$  ns) time resolution. Along with the DT and CSC systems, the RPC system provides the Level-1 trigger. It also capable to identify unambiguously the relevant bunch-crossing to which a muon track is associated even in the presence of the high rate and background at a full LHC luminosity.

Muon identification efficiency in the central region ( $|\eta| < 1$ ) is above 70% for muons with  $p_T > 5$  GeV/ $c$ . In the endcap regions ( $1 < |\eta| < 2.5$ ) identification efficiency is above 70% already for muons with  $p_T > 3$  GeV/ $c$ .

### 4.2.2 Tracker system

The CMS tracker system based only on silicon detectors (220 m<sup>2</sup> of Si): micro-strip and pixel. The strip detector consists of 10 barrel layers and 9 disks positioned both forward and backward. Depending on rapidity the high  $p_T$  tracks leaves 10 to 14 hits. The hit resolution is  $\sim 50$   $\mu$ m in

$r - \phi$  direction and  $\sim 500 \mu\text{m}$  in  $z$ -direction. The pixel system is placed closer to the interaction point and consists of 3 barrel layers and 2 disks positioned both forward and backward. Since pixel dimension is  $100 \times 150 \mu\text{m}^2$  the detector provides precise 2D information. The hit resolution is  $\sim 10 \mu\text{m}$  in  $r - \phi$  direction and  $\sim 17 \mu\text{m}$  in  $z$ -direction. Momentum resolution of the CMS tracker system varies from 0.5 % in the central region to 2 % in the endcaps for tracks with  $p_T = 1 \div 10 \text{ GeV}/c$ . The primary and secondary vertex resolution is event dependent. Usually, for  $b$ -decay channels the primary vertex resolution in the transverse plane is about  $20 \mu\text{m}$  and the secondary one is  $70 \div 100 \mu\text{m}$ .

### 4.3 Trigger strategies

Triggering in CMS is done in two steps. The Level-1 trigger is based on muon and calorimetry information. It has a latency of  $3.2 \mu\text{s}$  with a goal to reduce an event rate from 40 MHz to 100 kHz. At the second step called High-Level Trigger (HLT), information from all subsystems are readout and used in the event reconstruction. The reconstruction should be fast, therefore it is done locally: topologically around Level-1 pattern. At HLT the event rate decreases from 100 kHz to 100 Hz.

B-physics events are relatively soft. Hence, Level-1 calorimetry triggers having high  $E_T$  thresholds do not 'see' such events. Only events with one or two muons in the final state can be triggered with high selection efficiency for soft  $b\bar{b}$  events (see, for example, the Level-1 trigger menu in [44]). The transverse momentum thresholds of these triggers depend on an instantaneous luminosity, but will be kept as low as possible in the range  $p_T > 7 \div 14 \text{ GeV}/c$  for single and  $p_T > 3 \div 7 \text{ GeV}/c$  for di-muon triggers. At HLT, exclusive and inclusive  $b$ -triggers, based on partial reconstruction of searched  $b$ -decay channels, are used. For the final states with two muons, a selection procedure which is used the reconstructed di-muon secondary vertex significantly improves a signal over background ratio. The detailed description of the trigger algorithms can be found in [45].

### 4.4 Physics channels

The CMS heavy flavour menu could be subdivided into two categories. The first one is approved results, which will be reported further. The second category is not finished or not yet approved by the Collaboration active studies. The last category will be only mentioned below.

#### 4.4.1 Inclusive $b$ production

Three different mechanisms contribute to the heavy flavour production at hadron colliders: gluon splitting, gluon/quark fusion and flavour excitation. Each of these production mechanism has its own final state topology. It is important to measure the B-hadron  $p_T$  spectra within a large range to be able to disentangle the contributions of those mechanisms. In CMS the measurement of the inclusive  $b$  production cross section will be done with events containing jets and at least one muon.

The measurement of the differential cross sections is studied for B-hadrons of  $p_T > 50 \text{ GeV}/c$  and within the rapidity region of  $|\eta| < 2.4$ . The event selection requires a  $b$ -tagged jet to be present in the event. B tagging is based on inclusive secondary vertex reconstruction in

jets [46]. As the Level-1 trigger, the single muon one is used with the threshold of 19 GeV/ $c$ . At HLT in addition to the muon a  $b$ -jet with  $p_T > 50$  GeV/ $c$  is required.

The signal fraction is determined from a fit to the data distribution using the simulated shapes for the signal and background. Each reconstructed muon is associated to the most energetic  $b$ -tagged jet. The average efficiency of associating the muon with the  $b$ -tagged jet is 75%. The transverse momentum of the muon with respect to the  $b$ -jet axis discriminates signal against background.

Several sources of systematic uncertainties are considered in the study [47]. The largest uncertainty arises from the 3% error on a jet energy scale which leads to a cross section error of 12% at  $p_T > 50$  GeV/ $c$ .

1.6 million  $b$ -events for 1 fb $^{-1}$  of an integrated luminosity will be collected to investigate the  $b$  production mechanism in CMS. The  $b$  purity of the selected events varies as function of the transverse momentum in a range from 70% to 55%. The  $b$  production cross-section at  $p_T = 1.2$  TeV/ $c$  can be measured with 20% uncertainty.

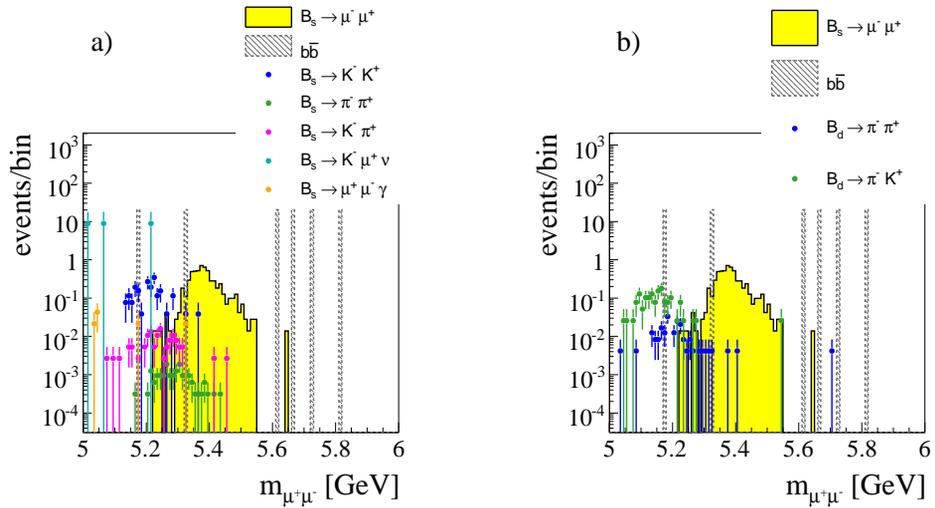


Fig. 21: Background  $m_{\mu\mu}$  distribution after the application of all selection criteria (with factorizing selection criteria) for all channels that are left: a)  $B_s$  decays, b)  $B_d$  decays.

#### 4.4.2 $B_s^0 \rightarrow \mu^+\mu^-$

Purely leptonic  $B$ -decays are theoretically very clean, thus providing an ideal environment for seeking indirect hints of NP effects. The SM branching ratio of  $B_s^0 \rightarrow \mu^+\mu^-$  is very small,  $(3.42 \pm 0.54) \cdot 10^{-9}$  [48], while in large- $\tan \beta$  NP models it can be enhanced by orders of magnitude [49]. Up to now only the upper limit on the branching ratio is set by the CDF Collaboration:  $4.7 \times 10^{-8}$  at the 90% C.L. [50].

The main challenge in the measurement of the  $B_s^0 \rightarrow \mu^+\mu^-$  decay rate is background suppression. Many background sources can mimic the signal topology. First, non-resonant  $b\bar{b}$  events

Table 3: Background for  $B_s^0 \rightarrow \mu^+ \mu^-$  samples used in the analysis. The visible cross-section, and the corresponding number of events for  $10 \text{ fb}^{-1}$  is given. The visible cross-sections include fragmentation, branching fractions, and (fake) muon  $p_T$  and  $|\eta|$  selection criteria. The numbers  $N_{\mu ID}$  do not yet include any selection criteria but hadron misidentification probability.

Sample	Generator cuts/channels	$\sigma_{vis}[\text{fb}]$	$N_{\mu ID} (10 \text{ fb}^{-1})$
$b\bar{b} \rightarrow \mu^+ \mu^- + X$	$p_T^\mu > 3 \text{ GeV}/c,  \eta^\mu  < 2.4$	$1.74 \times 10^7$	$1.74 \times 10^8$
	$p_T^{\mu\mu} > 5 \text{ GeV}/c$		
	$0.3 < \Delta R(\mu\mu) < 1.8$		
	$5 < m_{\mu\mu} < 6 \text{ GeV}/c^2$		
$B_s$ decays	$B_s \rightarrow K^- K^+$	$2.74 \times 10^5$	274
	$B_s \rightarrow \pi^- \pi^+$	$9.45 \times 10^3$	3
	$B_s \rightarrow K^- \pi^+$	$3.08 \times 10^4$	16
	$B_s \rightarrow K^- \mu^+ \nu$	$2.80 \times 10^5$	$2.80 \times 10^4$
	$B_s \rightarrow \mu^+ \mu^- \gamma$	$1.29 \times 10^1$	130
$B_d$ decays	$B_d \rightarrow \pi^- \pi^+$	$8.34 \times 10^4$	21
	$B_d \rightarrow \pi^- K^+$	$3.74 \times 10^5$	187
	$B_d \rightarrow \pi^- \mu^+ \nu$	$1.25 \times 10^6$	$6.25 \times 10^4$
	$B_d \rightarrow \mu^+ \mu^- \pi_0$	$3.77 \times 10^1$	377
$B_u$ decay	$B_u \rightarrow \mu^+ \mu^- \mu^+ \nu$	$2.24 \times 10^3$	$2.24 \times 10^4$
$B_c$ decays	$B_c \rightarrow \mu^+ \mu^- \mu^+ \nu$	$2.01 \times 10^1$	201
	$B_c \rightarrow J/\psi \mu^+ \nu$	$1.89 \times 10^3$	$1.89 \times 10^4$
$\Lambda_b$ decays	$\Lambda_b \rightarrow p \pi^-$	$4.22 \times 10^3$	1
	$\Lambda_b \rightarrow p K^-$	$8.45 \times 10^3$	1
QCD hadrons	$5 < M(hh) < 6 \text{ GeV}/c^2$	$2.24 \times 10^{11}$	$1.12 \times 10^8$

with each  $b \rightarrow \mu X$  decays. Second, non-resonant QCD events, where two high  $p_T$  hadrons are misidentified as muons. And finally, rare  $B_d$ ,  $B^+$ ,  $B_s$  and  $\Lambda_b$  decays, comprising hadronic, semileptonic, and radiative decays. Some of these decays constitute a resonant background, like  $B_s \rightarrow K^+ K^-$ , others have a continuum di-muon invariant mass distribution. Potentially, the resonant background is the most dangerous one. But the contribution from such backgrounds is negligible due to the excellent mass resolution provided by the CMS detector. Fig. 21 shows background  $m_{\mu\mu}$  distribution after the application of all selection criteria for all channels that are left.

The current background simulation does not include muon samples due to hadronic in-flight decays or punch-through. It has been estimated that this hadronic component will increase the background level by about 10% (see Fig. 5.10.9 in [51]). Table 3 summarizes studied so far background samples. The probabilities for hadron misidentification used to calculate expected number of background events in CMS are found to be  $\varepsilon_\pi = 0.5\%$ ,  $\varepsilon_K = 1.0\%$ ,  $\varepsilon_p = 0.1\%$ .

As the Level-1 trigger the di-muon one with a threshold of  $p_T > 3$  GeV/ $c$  for each muon is used. The HLT strategy relies critically on the tracker detector for a fast and high-efficiency reconstruction of the primary and secondary vertexes, the determination of muon momenta and the mass of the muon pair. Two muons are reconstructed with only 6 hits per a track in the tracker system and required to have transverse momentum of  $p_T > 4$  GeV/ $c$ , to be in the central part of the detector  $|\eta| < 2.4$  and to have opposite charges. Vertexing the two muons provides a powerful handle in the rate reduction: the quality of the vertex fit must be  $\chi^2 < 20$ . The three-dimensional flight length is required to be  $l_{3D} > 150$   $\mu\text{m}$ . The invariant mass of the muon pair is required to be in a tight window (150 MeV/ $c^2$ ) around the  $B_s$  meson mass.

For the offline analysis all tracks are reconstructed with full detector information. The same as above but tighter (e.g.  $\chi^2 < 1$ ) and additional selections are used to suppress background. The  $\eta\phi$  separation of the two muons  $\Delta R(\mu\mu) = \sqrt{\Delta\eta^2 + \Delta\phi^2}$  a powerful discriminator against gluon-gluon fusion background with both  $b$ -hadrons decaying semileptonically and must be in the range  $0.3 \div 1.2$ . The transverse momentum vector of the  $B_s$  candidate must be close to the displacement of the secondary vertex from the primary vertex: the cosine of the opening angle between the two vectors must fulfill  $\cos(\alpha) > 0.9985$ . The isolation  $I$  is determined from the  $B_s$  candidate transverse momentum and charged tracks with  $p_T > 0.9$  GeV/ $c$  in a cone with half-opening  $r = 1.0$  around the di-muon direction as follows:  $I = p_T^{\mu\mu} / (p_T^{\mu\mu} + \sum_{trk} |p_T|)$  and required to be  $I > 0.85$ . The significance of the flight length is required to be  $l_{3D}/\sigma_{3D} > 18.0$ . Mass separation between a di-muon candidate and the nominal  $B_s$  mass should not exceed 100 MeV/ $c^2$ .

For an integrated luminosity of  $10 \text{ fb}^{-1}$ , the expected number of signal events is  $n_S = 6.1 \pm 0.6_{stat} \pm 1.5_{sys}$ . The number of background events is  $n_B = (14.1)_{-14.1}^{+22.3}$ . An upper limit, extracted using the Bayesian approach, is  $Br(B_s^0 \rightarrow \mu^+\mu^-) < 1.4 \cdot 10^{-8}$  at the 90% CL [52].

#### 4.4.3 $B_s \rightarrow J/\psi\phi$

Important properties of  $B_s^0$  system can be studied with the decay channel  $B_s^0 \rightarrow J/\psi\phi$ , such as the width and mass difference of the two weak eigenstates:  $\Delta\Gamma_s, \Delta m_s$ . In addition, the decay  $B_s^0 \rightarrow J/\psi\phi$  is a golden channel for CP violation measurements. The particular spin structure of this decay allows to express its time dependence in a particular basis, called *transversity basis*, as the sum of 6 amplitudes where physical parameters enter differently. The most important of them is the weak phase  $\phi_s$  which is at present strongly constrained by CKM fits [53] and could represent a clear hint of NP if found to be significantly different from this prediction. At present the flavour tagging tools required to extract the weak phase are not yet available in CMS and only the mixing measurement is considered here.

The Level-1 trigger is based on di-muon selection with  $p_T > 3$  GeV/ $c$  each. HLT is the same as the one for  $B_s^0 \rightarrow \mu^+\mu^-$ , with additional requirement for the di-muon invariant mass to be within 150 MeV/ $c^2$  of the  $J/\psi$  mass. The  $J/\psi$  vertex is required to be  $3\sigma$  away from the primary vertex. The cosine of the angle between the transverse momentum vector and the transverse decay length vector of the  $J/\psi$  candidate is required to exceed 0.9.  $\phi$  candidates are reconstructed from all oppositely charged track pairs and all pairs with invariant mass within 20 MeV/ $c^2$  of the  $\phi$  mass are retained. All four tracks are then used for  $B_s^0$  candidates, requiring invariant mass within 200 MeV/ $c^2$  of nominal one. The transverse momentum of  $\phi$  and  $B_s^0$

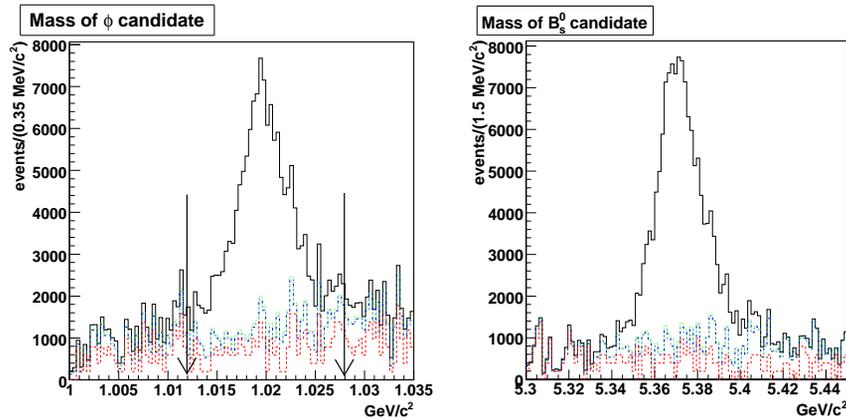


Fig. 22: Invariant mass of  $\phi$  (left) and  $B_s$  (right) candidates after all other cuts (except for the  $\phi$  mass requirement) have been applied; the selection on the  $\phi$  mass is indicated. Background is from inclusive  $b \rightarrow J/\psi X$  (red dashed line), from  $B_d^0 \rightarrow J/\psi K^{*0}$  (blue dashed-dotted line), and from combinatorial in signal events (green dotted line).

candidates are required to be greater than 1.0 and 5.0  $\text{GeV}/c$ , respectively.

The offline analysis follows the same criteria of HLT, but with complete information and tighter cuts. The main backgrounds arise from prompt  $J/\psi$  and  $B_d^0 \rightarrow J/\psi K^{*0}$  events. Fig. 22 shows invariant mass of  $\phi$  and  $B_s$  candidate distributions after all but  $\phi$  mass selections applied. From an untagged time-dependent analysis of  $B_s^0$  candidates the mixing parameters can be extracted. The result of the analysis shows that a first measurement of  $\Delta\Gamma_s/\Gamma_s$  can be made with 20% precision with an integrated luminosity of  $1.3 \text{ fb}^{-1}$ , while 5% precision can be reached with  $10 \text{ fb}^{-1}$  [54].

#### 4.4.4 $B_c^\pm \rightarrow J/\psi\pi^\pm$

The interest to the  $B_c$  meson relates to the uniqueness of the heavy-heavy quark system which carries flavour. The  $B_c$  meson has been observed at Tevatron and its mass and lifetime have been measured [55]. But available statistics does not allow to make such measurements precise enough and investigate properties of this system in details.

In CMS a feasibility study has been performed in the decay channel  $B_c^\pm \rightarrow J/\psi\pi^\pm$  [56]. First  $J/\psi$  candidates are composed by two muons with  $p_T > 4 \text{ GeV}/c$  and  $|\eta| < 2.2$  with opposite charge. The candidate invariant mass is required to be in the region from 3.0 to 3.2  $\text{GeV}/c^2$ . Then, pion candidates are selected by requiring a third track coming from the same vertex as the two muons with  $p_T > 2 \text{ GeV}/c$  and  $|\eta| < 2.4$ . The following selections are applied in addition: a proper decay length  $L_{xy}^{PDL} > 60 \mu\text{m}$ , a significance  $L_{xy}/\sigma_{xy} > 2.5$  and an opening angle between the vector from primary to secondary vertex and the momentum vector of the reconstructed  $B_c$ :  $\cos\theta > 0.8$ .

For  $1 \text{ fb}^{-1}$  120 signal and less than 3 background events are expected. The reconstructed  $B_c$  mass is  $6.4 \text{ GeV}/c^2$  and the width of the mass peak is  $15 \text{ MeV}/c^2$ . To extract the  $B_c$  meson

lifetime a binned likelihood fit was performed, resulting in  $\tau = 460 \pm 45$  fs.

#### 4.4.5 Additional heavy flavour decays

Although the following analysis are not yet finalized, it is worth while to list them here to demonstrate a future spectrum of heavy flavour activity in CMS.

Measurements of  $b\bar{b}$  production cross-section and lifetimes of B-mesons will be done in the following decay channels:  $B^\pm \rightarrow J/\psi K^\pm \rightarrow \mu^+\mu^- K^\pm$ ,  $B \rightarrow D^0 \mu X$ ,  $b \rightarrow J/\psi + X \rightarrow \mu^+\mu^- + X$ . The correlation study of  $J/\psi$  vs  $\mu$  provides clean measurements of  $b\bar{b}$  production mechanisms. Searches for NP is also planned to be done in  $B_s \rightarrow \mu^+\mu^-\gamma$ ,  $B \rightarrow (\phi, K^*, K_s)\mu^+\mu^-$  decay channels.

## 4.5 Conclusions

While designed for high- $p_T$  physics, the CMS has a broad heavy flavour program. Main features which allow this program are 1) high  $b\bar{b}$  event rate even at a low ( $10^{32} - 10^{33}$ ) initial luminosity, 2) the efficient low  $p_T$  di-muon trigger and 3) excellent tracking that provides high momentum, mass, vertex resolution. Expected results are competitive with current B-physics experiments.

#### Acknowledgments

One of us (A.S.) would like to thank V. Ovakimian for his help in the preparation of this paper.

## 5 Top production in the CMS experiment at the LHC

*Authors: C. Rosemann and R. Wolf*

### 5.1 Top quark pair production in CMS

The LHC will provide proton proton collisions at center-of-mass energies of 14 (10) TeV with a specific luminosity of  $\approx 10$  (0.1)  $\text{nb}^{-1}\text{s}^{-1}$ . This will allow the inclusive production of top anti-top quark pairs at a rate of 100-(10) Hz (where the values in parenthesis are given for startup). The cross section for the production of top anti-top quark pairs in proton proton collisions at these center-of-mass energies is expected to be  $908 \pm 83 \pm 30$  pb ( $414 \pm 40 \pm 20$  pb) [57], where the first error reflects scale uncertainties and the second error uncertainties in the choice and parametrization of parton density functions (PDFs). In the standard model (SM) top anti-top quark pair production is dominated by gluon gluon fusion (with a fraction of  $\approx 90\%$ ). As top quark production at the reached center-of-mass energies and luminosities at HERA is unaccessible the obvious impact of the HERA experiments on top production at the LHC lies in the determination of the most undefined gluon density function and thus in the reduction of the second uncertainty of the above cross section estimate. In figure 23 (left) three sets of different PDFs from the two HERA experiments H1 and ZEUS and from the CTEQ collaboration are shown at a scale of  $Q^2 = 10 \text{ GeV}^2$ . Due to kinematics and reconstruction requirements these PDFs will mostly be probed at large scales and medium to high proton momentum fraction  $x > 0.01$ , where these PDFs show smallest uncertainties and deviations.

In the beginning the emphazise will be put on the rediscovery of the top quark within the first (50-100)  $\text{pb}^{-1}$ , followed by inclusive cross section measurements based on robust selection

methods and first attempts to determine the top mass and differential cross sections with target luminosities of  $\approx 1 \text{ fb}^{-1}$ . An important aspect of top quark physics will be the capacity of top anti-top quark pair production of being a standard candle within the SM to exploit and demonstrate the detector understanding of the two major experiments ATLAS and CMS. The rediscovery of the top quark in early data taken with the CMS detector and prospects for first mass measurements in  $\approx 1 \text{ fb}^{-1}$  will be discussed in the following.

## 5.2 Rediscovery of the top quark

As an example for the rediscovery potential of the top quark within the first data a recent study in the semi-leptonic decay channel with a muon in the final state is presented [58]. It was performed for a target luminosity of  $10 \text{ pb}^{-1}$  of data at a center-of-mass energy of  $14 \text{ TeV}$  taken with the CMS detector. The conclusions may though be translated into equivalent conclusions for a luminosity of  $50 \text{ pb}^{-1}$  of data at a center-of-mass energy of  $10 \text{ TeV}$ . Focus was put on robust and simple selection methods with minimal dependency on detector components that may be least understood during startup. Main backgrounds are considered to be  $W$  or  $Z$  boson production with additional hard jets and QCD multijet production with leptons from  $b/c$  quark or in-flight decays which are mis-interpreted as originating from real  $W$  decays. Inclusive top anti-top pair production and  $W$  and  $Z$  boson production were produced with the Alpgen event generator for  $2 \rightarrow 4(5)$  processes in leading order and matched with parton showers using Pythia (values in parenthesis are given for inclusive top anti-top pair production). QCD multijet events were produced in leading order using Pythia. All events were passed through the full simulation of the

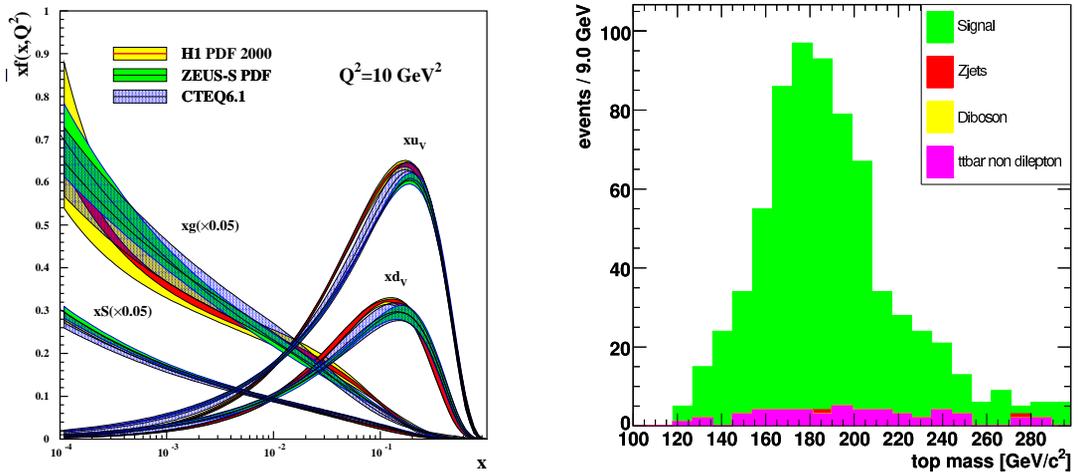


Fig. 23: (Left) Parton distribution functions (PDFs) from the HERA experiments H1 and ZEUS and from the CTEQ collaboration at a scale of  $Q^2 = 10 \text{ GeV}^2$ . Of relevance for top quark production at the LHC are the gluon density functions at values  $x > 0.01$  of the fractional proton momentum  $x$ . (Right) Top mass reconstructed from top anti-top quark pairs in the di-leptonic decay channel as expected from full simulation with the CMS detector with  $1 \text{ fb}^{-1}$  of integrated luminosity.

CMS detector including a simulation of the CMS L1 and High-Level-Trigger and mis-alignment and mis-calibration of the track detector and calorimeters as expected for the first  $10 \text{ pb}^{-1}$  were taken into account. Top events are selected by requiring a single isolated muon with transverse momentum  $p_T > 30 \text{ GeV}$  in the pseudo-rapidity range of  $|\eta| < 2.1$ , a leading jet with calibrated  $p_T > 65 \text{ GeV}$  and at least three further jets with calibrated  $p_T > 30 \text{ GeV}$  in a pseudo-rapidity range of  $|\eta| < 2.4$  in addition to the trigger criteria. With this selection the trigger efficiency is estimated to be above 90%. The muon is considered isolated if the summed  $p_T$  of all reconstructed tracks within a cone of  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.3$  in the vicinity of the muon does not exceed  $3 \text{ GeV}$  and the summed calorimeter entries in the same area do not exceed  $1 \text{ GeV}$ . Either  $b$ -tag information nor information on missing transverse energy are taken into consideration for the sake of a more robust event selection. In addition any of the selected jets is required to be separated from the isolated muon by at least 0.3 units in  $\Delta R$ .

This selection leads to a total of 128 events with top anti-top quark pairs in the semi-leptonic decay channel with a muon in the final state (with an estimated overall efficiency of 10%), 25 events with top anti-top pairs in other decay modes, 45  $W$ +jets events, 7  $Z$ +jets events and 11 multijet QCD events. The  $S/B$  is estimated to be 1.5 : 1 and the  $S/B(QCD)$  is estimated to be 11 : 1 with large uncertainties. As it is clear that the background from QCD multijet events will be the most difficult to control and to model methods for its estimation from data are discussed. This is a process still ongoing within the collaboration.

### 5.3 First measurements of the top mass

As an example for a measurement of the top quark mass within first data a study in the di-leptonic decay channel with  $1 \text{ fb}^{-1}$  of data taken with the CMS detector is presented [59]. Due to its clear event topology this channel is expected to have the best signal to background ratio. Main backgrounds are considered to be  $Z$  and di-boson production associated with additional hard jets and top events from other decay channels. All events were produced with the Pythia event generator in leading order and passed through the full simulation of the CMS detector including a simulation of the CMS L1 and High-Level-Trigger. Top events are selected by requiring two isolated leptons ( $e$  or  $\mu$ ) of opposite sign with transverse momentum  $p_T > 20 \text{ GeV}$ , two jets with  $p_T > 30 \text{ GeV}$  and missing transverse momentum larger than  $40 \text{ GeV}$ , in addition to the trigger criteria. Leptons are considered isolated if the summed  $p_T$  of all reconstructed tracks within a cone of  $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} < 0.2$  in the vicinity of the lepton does not exceed 10% of the lepton's momentum. Electrons are identified using a likelihood method exploiting shower shape characteristics and the matching of tracks and calorimeter objects. For leptons of the same type an additional veto on the  $Z$  invariant mass is implied. The jets are required to fulfill a  $b$ -tag requirement based on the  $p_T$  and the invariant mass of the associated tracks and the result of a combined  $b$ -tag algorithm [60]. This selection is expected to provide a signal over background ratio of 12 : 1 for a top mass estimate between 100 and 300 GeV the remaining background mostly originating from other top decay channels. In fig. 23 (right) the most likely top mass determined from a parameter scan in the range of (100-300) GeV is shown. Unknowns are reduced imposing constraints on momentum conservation in the transverse plane, the  $W$  invariant mass and the equality of the top mass in both decay branches. Remaining ambiguities are taken into account by a weighting procedure based on the SM expectation of the neutrino

momentum spectrum. The fit of a Gaussian function yields a top mass of  $m_{\text{top}}^{\text{rec}} = 178.5 \pm 1.5$  (stat.)  $\pm 4$  (syst.) GeV for an input mass of  $m_{\text{top}}^{\text{gen}} = 175$  GeV. Systematic uncertainties are expected to be dominated by the uncertainty of the validity of the imposed constraints in the presence of initial and final state radiation, and the uncertainty of the jet energy scale (JES). For  $10 \text{ fb}^{-1}$  the uncertainties are expected to be reduced to  $\Delta m_{\text{top}} = 0.5$  (stat.)  $\pm 1$  (syst.) GeV.

## References

- [1] Z. Lin and M. Gyulassy, Phys. Rev. **C51**, 2177 (1995).
- [2] Y. Dokshitzer and D. Kharzeev, Phys. Lett. **B519**, 199 (2001).
- [3] M. Djordjevic, M. Gyulassy, and S. Wicks, Phys. Rev. Lett. **94**, 112301 (2005).
- [4] PHENIX Collaboration, S. Adler *et al.*, Phys. Rev. Lett. **96**, 032301 (2006).
- [5] PHENIX Collaboration, A. Adare *et al.*, Phys. Rev. Lett. **98**, 172301 (2007).
- [6] STAR Collaboration, B. Abelev *et al.*, Phys. Rev. Lett. **98**, 192301 (2007).
- [7] R. Vogt, Eur. Phys. J. Special Topics **155**, 213 (2008).
- [8] STAR Collaboration, A. Mischke *et al.*, J. Phys. **G35**, 104117 (2008). ArXiv: 0804.4601.
- [9] A. Mischke, Phys. Lett. B. in press. ArXiv: 0807.1309.
- [10] M. Cacciari, P. Nason, and R. Vogt, Phys. Rev. Lett. **95**, 122001 (2005).
- [11] ALICE Collaboration, A. Dainese *et al.*, J. Phys. **G35**, 044046 (2008).
- [12] ALICE Collaboration, F. Carminati *et al.*, J. Phys. **G30**, 1517 (2005).
- [13] ALICE Collaboration, B. Alessandro *et al.*, J. Phys. **G32**, 1295 (2006).
- [14] Y. Dokshitzer and D. Kharzeev, Phys. Lett. **B519**, 199 (2001).
- [15] N. Armesto, A. Dainese, C. Salgado, and U. Wiedemann, Phys. Rev. **D71**, 054027 (2005).
- [16] M. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).
- [17] ALICE Collaboration, E. Bruna *et al.*, Preprint **nucl-ex/0703005** (2007).
- [18] S. Eidelmann and others [Particle Data Group Collaboration], Phys. Lett. **B592**, 1 (2004).
- [19] T. Sjöstrand, P. Edén, C. Friberg, L. Lönnblad, G. Miu, S. Mrenna, and E. Norrbin, Computer Phys. Commun. **135**, 238 (2001).
- [20] F. Antinori, A. Dainese, M. Lunardon, and R. Turrisi, ALICE Internal Note **033** (2006).
- [21] A. Dainese and M. Masera, ALICE Internal Note **027** (2003).
- [22] UA1 Collaboration, C. Albajar *et al.*, Phys. Lett. **B 213**, 405 (1988).

- [23] P. Crochet, R. Guernane, A. Morsch, and E. Vercellin, ALICE Internal Note **018** (2005).
- [24] J. Baines *et al.* hep-ph/0601164.
- [25] M. Cacciari, M. Greco, and P. Nason, JHEP **9805**, 007 (1998).
- [26] H. Jung, Computer Phys. Commun. **143**, 100 (2002).
- [27] M. Bedjidian *et al.* hep-ex/0311048.
- [28] CDF Collaboration, F. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005).  
hep-ex/0412071;  
CDF Collaboration, F. Acosta *et al.*, Phys. Rev. Lett. **88**, 161802 (2002).
- [29] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C4**, 463 (1998).  
hep-ph/9803445.
- [30] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Phys. Lett.  
**B531**, 216 (2002). hep-ph/0201127.
- [31] CTEQ Collaboration, H. L. Lai *et al.*, Eur. Phys. J. **C12**, 375 (2000). hep-ph/9903282.
- [32] J. Pumplin *et al.*, JHEP **07**, 012 (2002).
- [33] ATLAS Collaboration, ATLAS TDR 15, CERN/LHCC 99-15 (1999).
- [34] D. Emeliyanov *et al.*, ATL-DAQ-CONF-2007-031 (2007).
- [35] P. Jussel *et al.*, ATL-PHYS-PUB-2007-003 (2007).
- [36] S. George, ATL-DAQ-2004-004 (2004).
- [37] CDF Collaboration, F. Acosta *et al.*, Phys. Ref. **D71**, 032001 (2005).
- [38] W. Yao *et al.*, J. of Phys. G **33**, 1 (2006).
- [39] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **69**, 3704 (1992).
- [40] CDF Collaboration, F. Abulencia *et al.*, Phys. Rev. Lett. **99**, 132001 (2007).
- [41] G. Bodwin, E. Braaten, and G. Lepage, Phys. Rev. **D51**, 1125 (1995).
- [42] ATLAS Collaboration, CERN-OPEN-2008-020 (2008).
- [43] CMS Collaboration, S. Chatrchyan *et al.*, JINST **3**, S08004 (2008).
- [44] CMS Collaboration, CERN-LHCC-2006-021 (2006).
- [45] CMS Collaboration, CERN-LHCC-2002-026 (2006).
- [46] C. Weiser, CMS NOTE-2006/014 (2006).

- [47] V. Andreev, D. Cline, and S. Otwinowski, CMS NOTE-2006/120 (2006).
- [48] A. Buras, Phys. Lett. **B566**, 115 (2003).
- [49] K. Babu and C. Kolda, Phys. Rev. Lett. **84**, 228 (2000).
- [50] CDF Collaboration, T. Aaltonen *et al.*, Phys. Rev. Lett. **100**, 101802 (2008).
- [51] CMS Collaboration, J. Layter *et al.*, CERN-LHCC-1997-032 (1997).
- [52] U. Langenegger, CMS-CR-2006/071 (2006).
- [53] M. Bona *et al.*, JHEP **0610**, 081 (2006).
- [54] V. Ciulli *et al.*, CMS NOTE-2006/121 (2006).
- [55] CDF Collaboration, Phys. Rev. Lett. **96**, 082002 (2006).
- [56] X. Meng, J. Tao, and G. Chen, CMS NOTE-2006/118 (2006).
- [57] M. Cacciari *et al.*, arXiv:0804.2008.
- [58] CMS Collaboration, CMS PAS Top08-005.
- [59] CMS Collaboration, J. Phys. **G34**, 995 (2007).
- [60] R. Kinnunen and S. Lehti, CMS NOTE 2006/075.

# Heavy flavour production at the LHC: Theoretical Aspects

Z. J. Ajaltouni<sup>a</sup>, A. Banfi<sup>b</sup>, S. Baranov<sup>c</sup>, I. Bierenbaum<sup>d</sup>, J. Blümlein<sup>e</sup>, G. Corcella<sup>f</sup>,  
M. Czakon<sup>g</sup>, G. Ferrera<sup>h</sup>, S. Klein<sup>e</sup>, B.A. Kniehl<sup>i</sup>, G. Kramer<sup>j</sup>, A. Likhoded<sup>j</sup>, D.A. Milstead<sup>k</sup>,  
O.I. Piskounova<sup>l</sup>, V.A. Saleev<sup>m</sup>, I. Schienbein<sup>n</sup>, H. Spiesberger<sup>o</sup>, R.S. Thorne<sup>p</sup>, W.K. Tung<sup>q</sup>,  
G. Zanderighi<sup>r</sup>, N. Zotov<sup>s</sup>

<sup>a</sup> LPC, Université Blaise Pascal, F-63177 Aubière Cedex, France

<sup>b</sup> INFN, Sezione di Milano-Bicocca, Italy

<sup>c</sup> P.N. Lebedev Institute of Physics, Lenin avenue 53, Moscow 119991, Russia

<sup>d</sup> Deutsches Elektronen-Synchrotron, Platanenallee 6, D-15738 Zeuthen, Germany (now at IFIC, Universitat de València, Spain)

<sup>e</sup> Deutsches Elektronen-Synchrotron, Platanenallee 6, D-15738 Zeuthen, Germany

<sup>f</sup> Museo Storico della Fisica e Centro Studi e Ricerche ‘E. Fermi’, Roma, and Scuola Normale Superiore and INFN Sezione di Pisa, Italy

<sup>g</sup> University Würzburg, Germany

<sup>h</sup> University and INFN, Firenze, Italy

<sup>i</sup> II. Inst. für Theoretische Physik, Universität Hamburg, Germany

<sup>j</sup> Theoretical Department, Institute for High Energy Physics (IHEP), Protvino, Moscow reg., Russia

<sup>k</sup> Stockholm University, Sweden

<sup>l</sup> Lebedev Institute, Moscow, Russia

<sup>m</sup> Samara State University, Samara, Russia

<sup>n</sup> LPSC, Grenoble, France

<sup>o</sup> University Mainz, Germany

<sup>p</sup> University College London

<sup>q</sup> Michigan State U. and U. Washington, Seattle

<sup>r</sup> The Rudolf Peierls Centre for Theoretical Physics, 1 Keble Road, University of Oxford, UK

<sup>s</sup> Institute of Nuclear Physics, Moscow State University, Lenin Hills, Moscow 119991, Russia

## Abstract

A proper inclusion of heavy quark mass effects in Parton Distribution Function fits has proved crucial. We present a review these effects in DIS and their impact on global analyses and lay out all elements of a properly defined general mass variable flavor number scheme (GM VFNS) that are shared by all modern formulations of the problem. We also report about progress in a number of theoretical problems related to exclusive measurements of heavy flavors. These topics include fragmentation functions for charmed mesons including finite mass effects, fragmentation functions including non-perturbative corrections based on an effective QCD coupling, a discussion of the status of higher-order calculations for top quark production and for polarized structure functions, heavy quark and quarkonium production in the Regge limit, double heavy baryon production, tests of time reversal and CP symmetry in  $\Lambda_b$  decays, as well as a study of the general properties of massive

exotic hadrons that will be relevant for an understanding of their detection at the LHC.

*Coordinators: M. Cacciari and H. Spiesberger*

## 1 PQCD Formulations with Heavy Quark Masses and Global Analysis

*Authors: R.S. Thorne and W.K. Tung*

### 1.1 Introduction

The proper treatment of heavy flavours in global QCD analysis of parton distribution functions (PDFs) is essential for precision measurements at hadron colliders. Recent studies [1–5] show that the standard-candle cross sections for  $W/Z$  production at the LHC are sensitive to detailed features of PDFs that depend on heavy quark mass effects; and certain standard model as well as beyond standard model processes depend crucially on better knowledge of the  $c$ -quark parton density, in addition to the light parton flavors. These studies also make it clear that the consistent treatment of heavy flavours in perturbative QCD (PQCD) require theoretical considerations that go beyond the familiar textbook parton picture based on massless quarks and gluons. There are various choices, explicit and implicit, which need to be made in various stages of a proper calculation in generalised PQCD including heavy quark mass effects. In the global analysis of PDFs, these choices can affect the resulting parton distributions. Consistent choices are imperative; mistakes may result in differences that are similar to, or even greater than, the quoted uncertainties due to other sources (such as the propagation of input experimental errors). In this report, we will provide a brief, but full, review of issues related to the treatment of heavy quark masses in PQCD, embodied in the general mass variable flavor scheme (GM VFNS).

In Sec. 1.2, we describe the basic features of the modern PQCD formalism incorporating heavy quark masses. In Sec. 1.3, we first delineate the common features of GM VFNS, then identify the different (but self-consistent) choices that have been made in recent global analysis work, and compare their results. For readers interested in practical issues relating to the use (or choice) of PDFs in various physics applications, we present a series of comments in Sec. 1.4 intended as guidelines. In Sec. 1.5, we discuss the possibility of intrinsic heavy flavors.

We note that, this review on GM VFNS and global analysis is not intended to address the specific issues pertinent to heavy flavor production (especially the final state distributions). For this particular process, somewhat different considerations may favor the adoption of appropriate fixed flavor number schemes (FFNS). We shall not go into details of these considerations; but will mention the FFNS along the way, since the GM VFNS is built on a series of FFNS's. We will comment on this intimate relationship whenever appropriate.

### 1.2 General Considerations on PQCD with Heavy Flavor Quarks

The quark-parton picture is based on the factorization theorem of PQCD. The conventional proof of the factorization theorem proceeds from the zero-mass limit for all the partons—a good approximation at energy scales (generically designated by  $Q$ ) far above all quark mass thresholds

(designated by  $m_i$ ). This clearly does not hold when  $Q/m_i$  is of order 1.<sup>1</sup> It has been recognised since the mid-1980's that a consistent treatment of heavy quarks in PQCD over the full energy range from  $Q \lesssim m_i$  to  $Q \gg m_i$  can be formulated [6]. In 1998, Collins gave a general proof of the factorization theorem (order-by-order to all orders of perturbation theory) that is valid for non-zero quark masses [7]. The resulting general theoretical framework is conceptually simple: it represents a straightforward generalisation of the conventional zero-mass (ZM) modified minimal subtraction ( $\overline{\text{MS}}$ ) formalism and it contains the conventional approaches as special cases in their respective regions of applicability; thus, it provides a good basis for our discussions.

The implementation of any PQCD calculation on physical cross sections requires attention to a number of details, both kinematical and dynamical, that can affect both the reliability of the predictions. Physical considerations are important to ensure that the right choices are made between perturbatively equivalent alternatives that may produce noticeable differences in practical applications. It is important to make these considerations explicit, in order to make sense of the comparison between different calculations in the literature. This is what we shall do in this section. In subsequent sections, we shall point out the different choices that have been made in recent global analysis efforts.

Heavy quark physics at HERA involve mostly charm ( $c$ ) and bottom ( $b$ ) production; at LHC, top ( $t$ ) production, in addition, is of interest. For simplicity, we often focus the discussion of the theoretical issues on the production of a single heavy quark flavor, which we shall denote generically as  $H$ , with mass  $m_H$ . The considerations apply to all three cases,  $H = c, b, \& t$ . For global analysis, the most important process that requires precision calculation is DIS; hence, for physical predictions, we will explicitly discuss the total inclusive and semi-inclusive structure functions, generically referred to as  $F^\lambda(x, Q)$ , where  $\lambda$  represents either the conventional label (1, 2, 3) or the alternative ( $T, L, 3$ ) where  $T/L$  stands for transverse/longitudinal respectively.

### 1.2.1 The Factorization Formula

The PQCD factorization theorem for the DIS structure functions has the general form

$$F_\lambda(x, Q^2) = \sum_k f_k \otimes C_k^\lambda = \sum_k \int_x^1 \frac{d\xi}{\xi} f_k(\xi, \mu) C_k^\lambda \left( \frac{\chi}{\xi}, \frac{Q}{\mu}, \frac{m_i}{\mu}, \alpha_s(\mu) \right). \quad (1)$$

Here, the summation is over the active parton flavor label  $k$ ,  $f^k(x, \mu)$  are the parton distributions at the factorization scale  $\mu$ ,  $C_k^\lambda$  are the Wilson coefficients (or hard-scattering amplitudes) that can be calculated order-by-order in perturbation theory. The lower limit of the convolution integral  $\chi$  is determined by final-state phase-space constraints: in the conventional ZM parton formalism it is simply  $x = Q^2/2q \cdot p$ —the Bjorken  $x$ —but this is no longer true when heavy flavor particles are produced in the final state, cf. Sec. 1.2.4 below. The renormalization and factorization scales are jointly represented by  $\mu$ : in most applications, it is convenient to choose  $\mu = Q$ ; but there are circumstances in which a different choice becomes useful.

---

<sup>1</sup>Heavy quarks, by definition, have  $m_i \gg \Lambda_{QCD}$ . Hence we always assume  $Q, m_i \gg \Lambda_{QCD}$ . In practice,  $i = c, b, t$ .

### 1.2.2 Partons and Schemes for General Mass PQCD

In PQCD, the summation  $\sum_k$  over “parton flavor” label  $k$  in the factorization formula, Eq. (1), is determined by the *factorization scheme* chosen to *define* the Parton Distributions  $f_k(x, \mu)$ .

If mass effects of a heavy quark  $H$  are to be taken into account, the simplest scheme to adopt is the *fixed flavor number scheme* (FFNS) in which all quark flavors below  $H$  are treated as zero-mass and one sums over  $k = g, u, \bar{u}, d, \bar{d}, \dots$  up to  $n_f$  flavors of *light* (massless) quarks. The mass of  $H$ ,  $m_H$ , appears explicitly in the Wilson coefficients  $\{C_k^\lambda\}$ , as indicated in Eq. 1. For  $H = \{c, b, t\}$ ,  $n_f = \{3, 4, 5\}$  respectively. Historically, higher-order ( $\mathcal{O}(\alpha_s^2)$ ) calculations of the heavy quark production [8] were all done first in the FFNS. These calculations provide much improved results when  $\mu(Q)$  is of the order of  $m_H$  (both above and below), over those of the conventional ZM ones (corresponding to setting  $m_H = 0$ ).

Unfortunately, at any finite order in perturbative calculation, the  $n_f$ -FFNS results become increasingly unreliable as  $Q$  becomes large compared to  $m_H$ : the Wilson coefficients contain logarithmic terms of the form  $\alpha_s^n \ln^m(Q/m_H)$ , where  $m = 1 \dots n$ , at order  $n$  of the perturbative expansion, implying they are not infrared safe—higher order terms do not diminish in size compared to lower order ones—the perturbative expansion eventually breaks down. Thus, even if all  $n_f$ -flavor FFNS are mathematically equivalent, in practice, the 3-flavor scheme yields the most reliable results in the region  $Q \lesssim m_c$ , the 4-flavor scheme in  $m_c \lesssim Q \lesssim m_b$ , the 5-flavor scheme in  $m_b \lesssim Q \lesssim m_t$ , and, if needed, the 6-flavor scheme in  $m_t \lesssim Q$ . (Cf. related discussions later in this section.)

This leads naturally to the definition of the more general *variable flavor number scheme* (VFNS): it is a *composite scheme* consisting of the sequence of  $n_f$ -flavor FFNS, each in its region of validity, for  $n_f = 3, 4, \dots$  as described above; and the various  $n_f$ -flavor schemes are related to each other by perturbatively calculable transformation (finite-renormalization) matrices among the (running) coupling  $\alpha_s$ , the running masses  $\{m_H\}$ , the parton distribution functions  $\{f_k\}$ , and the Wilson coefficients  $\{C_k^\lambda\}$ . These relations ensure that there are only one set of independent renormalization constants, hence make the definition of the composite scheme precise for all energy scale  $\mu(Q)$ ; and they ensure that physical predictions are well-defined and continuous as the energy scale traverses each of the overlapping regions  $Q \sim m_H$  where both the  $n_f$ -flavor and the  $(n_f + 1)$ -flavor schemes are applicable. The theoretical foundation for this intuitively obvious scheme can be found in [6, 7], and it was first applied in detail for structure functions in [9]. Most recent work on heavy quark physics adopt this general picture, in one form or another. We shall mention some common features of this general-mass (GM) VFNS in the next few paragraphs; and defer the specifics on the implementation of this scheme, as well as the variations in the implementation allowed by the general framework until Sec. 1.3.

As mentioned above, the  $n_f$ -flavor and the  $(n_f + 1)$ -flavor schemes within the GM VFNS should be matched at some *match point*  $\mu_M$  that is of the order of  $m_H$ . In practice, the matching is commonly chosen to be exactly  $\mu_M = m_H$ , since it has been known that, in the calculational scheme appropriate for GM VFNS<sup>2</sup>, the transformation matrices vanish at this particular scale

<sup>2</sup>Technically, this means employing the CWZ subtraction scheme [10] in calculating the higher-order Feynman diagrams. CWZ subtraction is an elegant extension of the  $\overline{\text{MS}}$  subtraction scheme that ensures the decoupling of heavy quarks at high energy scales order-by-order. This is essential for factorization to be valid at each order of perturbation theory. (In the original  $\overline{\text{MS}}$  subtraction scheme, decoupling is satisfied only for the full perturbation

at NLO in the perturbative expansion [6]; thus discontinuities of the renormalized quantities are always of higher order, making practical calculations simpler in general.

Strictly speaking, once the component  $n_f$ -flavor schemes are unambiguously matched, one can still choose an independent *transition scale*,  $\mu_T$ , at which to switch from the  $n_f$ -flavor scheme to the  $(n_f + 1)$ -flavor scheme in the calculation of physical quantities in defining the GM VFNS. This scale must again be within the overlapping region, but can be different from  $\mu_M$  [1, 7]. In fact, it is commonly known that, from the physics point of view, in the region above the  $m_H$  threshold, up to  $\eta m_H$  with a reasonable-sized constant factor  $\eta$ , the most natural parton picture is that of  $n_f$ -flavor, rather than  $(n_f + 1)$ -flavor one.<sup>3</sup> For instance, the 3-flavor scheme calculation has been favored by most HERA work on charm and bottom quark production, even if the HERA DIS kinematic region mostly involves  $Q > m_c$ ; and it is also used in the dynamically generated parton approach to global analysis [11].

In practice, almost all implementations of the GM VFNS simply choose  $\mu_T = \mu_M = m_H$  (often not explicitly mentioning the conceptual distinction between  $\mu_T$  and  $\mu_M = m_H$ ). The self-consistency of the GM VFNS guarantees that physical predictions are rather insensitive to the choice of the transition point as long as it is within the overlapping region of validity of the  $n_f$ - and  $(n_f + 1)$ -flavor ones. The simple choice of  $\mu_T = m_H$  corresponds to opting for the lower end of this region for the convenience in implementation. In the following, we shall use the terms matching point and transition point interchangeably. As with all definition ambiguities in perturbative theory, the sensitivity to the choice of matching and transition points diminishes at higher orders.

### 1.2.3 Treatment of Final-state Flavors

For total inclusive structure functions, the factorization formula, Eq. (1), contains an implicit summation over all possible quark flavors in the final state. One can write,

$$C_k = \sum_j C_k^j \quad (2)$$

where “ $j$ ” denotes final state flavors, and  $\{C_k^j\}$  represent the Wilson coefficients (hard cross sections) for an incoming parton “ $k$ ” to produce a final state containing flavor “ $j$ ” calculable perturbatively from the relevant Feynman diagrams. It is important to emphasize that “ $j$ ” labels quark flavors that can be produced *physically* in the final state; it is *not* a *parton* label in the sense of initial-state parton flavors described in the previous subsection. The latter (labeled  $k$ ) is a theoretical construct and scheme-dependent (e.g. it is fixed at three for the 3-flavor scheme); whereas the final-state sum (over  $j$ ) is over *all flavors* that can be physically produced. Furthermore, the initial state parton “ $k$ ” does not have to be on the mass-shell, and is commonly treated as massless; whereas the final state particles “ $j$ ” should certainly be *on-mass-shell* in order to satisfy the correct kinematic constraints for the final state phase space and yield physically meaningful

---

series—to infinite orders.)

<sup>3</sup>Specifically, the  $n_f$ -flavor scheme should fail when  $\alpha_s(\mu) \ln(\mu/m_H) = \alpha_s(\mu) \ln(\eta)$  ceases to be a small parameter for the effective perturbation expansion. However, no theory can tell us precisely how small is acceptably “small”—hence how large  $\eta$  is permitted. Ardent FFNS advocates believe even the range of the 3-flavor scheme extends to all currently available energies, including HERA [11]. For GM VFNS, see the next paragraph.

results.<sup>4</sup> Thus, in implementing the summation over final states, the most relevant physical scale is  $W$ —the CM energy of the virtual Compton process—in contrast to the scale  $Q$  that controls the initial state summation over parton flavors.

The distinction between the two summations is absent in the simplest implementation of the conventional (i.e., textbook) zero-mass parton formalism: if all quark masses are set to zero to begin with, then all flavors can be produced in the final state. This distinction becomes blurred in the commonly used zero-mass (ZM) VFNS, where the heavy quark masses  $\{m_H\}$  implicitly enter because the number of effective parton flavors is incremented as the scale parameter  $\mu$  crosses each heavy quark threshold. This creates apparent paradoxes in the implementation of the ZM VFNS, such as: for  $\mu = Q < m_b$ ,  $b$  is not counted as a parton, the partonic process  $\gamma + g \rightarrow b\bar{b}$  would not be included in DIS calculations, yet physically this can be significant if  $W \gg 2m_b$  (small  $x$ ); whereas for  $\mu = Q > m_b$ ,  $b$  is counted as a massless parton, the contribution of  $\gamma + g \rightarrow b\bar{b}$  to DIS would be the same as that of  $\gamma + g \rightarrow d\bar{d}$ , but physically this is wrong for moderate values of  $W$ , and furthermore, it should be zero if  $W < 2m_b$  (corresponding to large  $x$ ). (We shall return to this topic in Sec. 1.3.1.)

These problems were certainly overlooked in conventional global analyses from its inception until the time when issues on mass-effects in PQCD were brought to the fore after the mid 1990's [9, 12–15]. Since then, despite its shortcomings the standard ZM VFNS continues to be used widely because of its simplicity and because NLO Wilson coefficients for most physical processes are still only available in the ZM VFNS. Most groups produce the standard ZM VFNS as either their default set or as one of the options, and they form the most common basis for comparison between groups, e.g. the “benchmark study” in [16].

It is obvious that, in a proper implementation of PQCD with mass (in any scheme), the distinction between the initial-state and final-state summation must be unambiguously, and correctly, observed. For instance, even in the 3-flavor regime (when  $c$  and  $b$  quarks are *not counted as partons*), the charm and bottom flavors still need to be counted in the final state—at tree-level via  $W^+ + d/s \rightarrow c$ , and at 1-loop level via the gluon-fusion processes such as  $W^+ + g \rightarrow \bar{s} + c$  or  $\gamma + g \rightarrow c\bar{c}$  ( $b\bar{b}$ ), provided there is enough CM energy to produce these particles.

#### 1.2.4 Phase-space Constraints and Rescaling

The above discussion points to the importance of the proper treatment of final state phase space in heavy quark calculations. Once mass effects are taken into account, kinematic constraints have a significant impact on the numerical results of the calculation; in fact, they represent the dominant factor in the threshold regions of the phase space. In DIS, with heavy flavor produced in the final state, the simplest kinematic constraint that comes to mind is

$$W - M_N > \sum_f M_f \quad (3)$$

where  $W$  is the CM energy of the vector-boson–nucleon scattering process,  $M_N$  is the nucleon mass, and the right-hand side is the sum of *all* masses in the final state.  $W$  is related to the famil-

---

<sup>4</sup>Strict kinematics would require putting the produced heavy flavor mesons or baryons on the mass shell. In the PQCD formalism, we adopt the approximation of using on-shell final state heavy quarks in the underlying partonic process.

iar kinematic variables  $(x, Q)$  by  $W^2 - M_N^2 = Q^2(1-x)/x$ , and this constraint should ideally be imposed on the right-hand side of Eq. (1). Any approach achieving this represents an improvement over the conventional ZM scheme calculations, that ignores the kinematic constraint Eq. (3) (resulting in a gross over-estimate of the corresponding cross sections). The implementation of the constraint in the most usual case of NC processes, say  $\gamma/Z + c \rightarrow c$  (or any other heavy quark) is not automatic (and is absent in some earlier definitions of a GM VFNS) because in this partonic process one must account for the existence of a *hidden heavy particle*—the  $\bar{c}$ —in the target fragment. The key observation is, heavy objects buried in the target fragment are still a part of the final state, and should be included in the phase space constraint, Eq. (3).

Early attempts to address this issue were either approximate or rather cumbersome, and could not be naturally extended to high orders.<sup>5</sup> A much better physically motivated approach is based on the idea of rescaling. The simplest example is given by charm production in the LO CC process  $W + s \rightarrow c$ . It is well-known that, when the final state charm quark is put on the mass shell, kinematics requires the momentum fraction variable for the incoming strange parton,  $\chi$  in Eq. (1) to be  $\chi = x(1 + m_c^2/Q^2)$  [17], rather than the Bjorken  $x$ . This is commonly called the *rescaling variable*. The generalization of this idea to the more prevalent case of NC processes took a long time to emerge [18, 19] which extended the simple rescaling to the more general case of  $\gamma/Z + c \rightarrow c + X$ , where  $X$  contains only light particles, it was proposed that the convolution integral in Eq. (1) should be over the momentum fraction range  $\chi_c < \xi < 1$ , where

$$\chi_c = x \left( 1 + \frac{4m_c^2}{Q^2} \right) . \quad (4)$$

In the most general case where there are any number of heavy particles in the final state, the corresponding variable is (cf. Eq. (3))

$$\chi = x \left( 1 + \frac{(\sum_f M_f)^2}{Q^2} \right) . \quad (5)$$

This rescaling prescription has been referred to as ACOT $\chi$  in the recent literature [18–20].

Rescaling shifts the momentum variable in the parton distribution function  $f^k(\xi, \mu)$  in Eq. (1) to a higher value than in the zero-mass case. For instance, at LO, the structure functions  $F_\lambda(x, Q)$  are given by some linear combination of  $f^k(x, Q)$  in the ZM formalism; but, with ACOT $\chi$  rescaling, this becomes  $f^k(\chi_c, Q)$ . In the region where  $(\sum_f M_f)^2/Q^2$  is not too small, especially when  $f(\xi, \mu)$  is a steep function of  $\xi$ , this rescaling can substantially change the numerical result of the calculation. It is straightforward to show that, when one approaches a given threshold  $(M_N + \sum_f M_f)$  from above, the corresponding rescaling variable  $\chi \rightarrow 1$ . Since generally  $f^k(\xi, \mu) \rightarrow 0$  as  $\xi \rightarrow 1$ , rescaling ensures a smoothly vanishing threshold behavior for the contribution of the heavy quark production term to all structure functions. This results in a universal<sup>6</sup>, and intuitively physical, realization of the threshold kinematic constraint for all heavy flavor production processes that is applicable to all orders of perturbation theory. For this reason, most recent global analysis efforts choose this method.

<sup>5</sup>In [9], the threshold violation was minimized by an artificial choice of the factorization scale  $\mu(m_H, Q)$ . In [14, 15] the kinematic limit was enforced exactly by requiring continuity of the slope of structure functions across the matching point, resulting in a rather complicated expression for the coefficient functions in Eq.(1).

<sup>6</sup>Since it is imposed on the (universal) parton distribution function part of the factorization formula.

### 1.2.5 Difference between $\{F_\lambda^{\text{tot}}\}$ and $\{F_\lambda^{\text{H}}\}$ Structure Functions

In PQCD, the most reliable calculations are those involving infra-red safe quantities—these are free from logarithmic factors that can become large (thereby spoiling the perturbative expansion). The total inclusive structure functions  $\{F_\lambda^{\text{tot}}\}$  defined in the GM VFNS are infrared safe, as suggested by the discussion of Sec. 1.2.2 and proven in Ref. [7].

Experimentally, the semi-inclusive DIS structure functions for producing a heavy flavor particle in the final state is also of interest. Theoretically, it is useful to note that the structure functions  $\{F_\lambda^{\text{H}}\}$  for producing heavy flavor  $H$  are not as well defined as  $F_\lambda^{\text{tot}}$ .<sup>7</sup> To see this, consider the relation between the two,

$$F_\lambda^{\text{tot}} = F_\lambda^{\text{light}} + F_\lambda^{\text{H}}, \quad (6)$$

where  $F_\lambda^{\text{light}}$  denotes the sum of terms with only light quarks in the final state, and  $F_\lambda^{\text{H}}$  consist of terms with at least one heavy quark  $H$  in the final state. Unfortunately,  $F_\lambda^{\text{H}}(x, Q, m_H)$  is, strictly speaking, *not infrared safe* beyond order  $\alpha_s$  (1-loop): they contain residual  $\ln^n(Q/m_H)$  terms at higher orders (2-loop and up). The same terms occur in  $F_\lambda^{\text{light}}$  due to contributions from virtual  $H$  loops, with the opposite sign. Only the sum of the two, i.e. the total inclusive quantities  $F_\lambda^{\text{tot}}$  are infra-red safe. This problem could be addressed properly by adopting a physically motivated, infrared-safe cut-off on the invariant mass of the heavy quark pair, corresponding to some experimental threshold [21] in the definition of  $F_\lambda^{\text{H}}$  (drawing on similar practises in jet physics). In practice, up to order  $\alpha_s^2$ , the result is numerically rather insensitive to this, and different groups adopt a variety of less sophisticated procedures, e.g. including contributions with virtual  $H$  loops within the definition of  $F_\lambda^{\text{H}}$ . Nonetheless, it is prudent to be aware that the theoretical predictions on  $F_\lambda^{\text{H}}$  are intrinsically less robust than those for  $F_\lambda^{\text{tot}}$  when comparing experimental results with theory calculations.

### 1.2.6 Conventions for “LO”, “NLO”, ... calculations

It is also useful to point out that, in PQCD, the use of familiar terms such as LO, NLO, ... is often ambiguous, depending on which type of physical quantities are under consideration, and on the convention used by the authors. This can be a source of considerable confusion when one compares the calculations of  $F_\lambda^{\text{tot}}$  and  $F_\lambda^{\text{H}}$  by different groups (cf. next section).

One common convention is to refer LO results as those derived from tree diagrams; NLO those from 1-loop calculations, ... and so on. This convention is widely used; and it is also the one used in the CTEQ papers. Another possible convention is to refer to LO results as the *first non-zero term* in the perturbative expansion; NLO as one order higher in  $\alpha_s$ , ... and so on. This convention originated in FFNS calculations of heavy quark production; and it is also used by the MRST/MSTW authors. It is a process-dependent convention, and it depends *a priori* on the knowledge of results of the calculation to the first couple of orders in  $\alpha_s$ .

---

<sup>7</sup>In the following discussion, we shall overlook logarithmic factors normally associated with fragmentation functions for simplicity. These are similar to those associated with parton distributions, but are less understood from the theoretical point of view—e.g. the general proof of factorization theorem (with mass) [7] has not yet been extended to cover fragmentation.

Whereas the two conventions coincide for quantities such as  $F_2^{\text{tot}}$ , they lead to different designations for the longitudinal structure function  $F_L^{\text{tot}}$  and the  $n_f$ -flavor  $F_2^{\text{H},n_f}$ , since the tree-level results are zero for these quantities. These designations, by themselves, are only a matter of terminology. However, mixing the two distinct terminologies in comparing results of different groups can be truly confusing. This will become obvious later.

### 1.3 Implementations of VFNS: Common Features and Differences

In this section, we provide some details of the PQCD basis for the GM VFNS, and comment on the different choices that have been made in the various versions of this general framework, implemented by two of the major groups performing global QCD analysis.

#### 1.3.1 Alternative Formulations of the ZM VFNS

As pointed out in Sec. 1.2.3, the ZM VFNS, as commonly implemented, represents an unreliable approximation to the correct PQCD in some kinematic regions because of inappropriate handling of the final-state counting and phase-space treatment, in addition to the neglect of heavy-quark mass terms in the Wilson coefficients. Whereas the latter is unavoidable to some extent, because the massive Wilson coefficients have not yet been calculated even at 1-loop level for most physical processes (except for DIS), the former (which can be more significant numerically in certain parts of phase space) can potentially be remedied by properly counting the final states and using the rescaling variables, as discussed in Secs. 1.2.3 and 1.2.4 under general considerations. Thus, alternative formulations of the ZM VFNS are possible that only involve the zero-mass approximation in the Wilson coefficient. This possibility has not yet been explicitly explored.

#### 1.3.2 Parton Distribution Functions in VFNS (ZM and GM)

In PQCD, the factorization scheme is determined by the choices made in defining the parton distribution functions (as renormalized Green functions). In a GM VFNS based on the generalized  $\overline{\text{MS}}$  subtraction (cf. footnote 2) the evolution kernel of the DGLAP equation is *mass-independent*; thus the PDFs, so defined, apply to GM VFNS calculations as they do for the ZM VFNS.

In the VFNS, the PDFs switch from the  $n_f$ -flavor FFNS ones to the  $(n_f + 1)$ -flavor FFNS ones at the matching point  $\mu = m_H$  (cf. Sec. 1.2.2); the PDFs above/below the matching point are related, order-by-order in  $\alpha_s$ , by:

$$f_j^{VF}(\mu \rightarrow m_H^+) \equiv f_j^{(n_f+1)FF} = A_{jk} \otimes f_k^{n_f FF} \equiv A_{jk} \otimes f_j^{VF}(\mu \rightarrow m_H^-), \quad (7)$$

where  $m_H^{+/-}$  indicate that the  $\mu \rightarrow m_H$  limit is taken from above/below, and we have used the shorthand VF/FF for VFNS/FFNS in the superscripts. The transition matrix elements  $A_{jk}(\mu/m_H)$ , representing a finite-renormalization between the two overlapping FFNS schemes, can be calculated order by order in  $\alpha_s$ ; they are known to NNLO, i.e.  $\mathcal{O}(\alpha_s^2)$  [12, 13]. (Note that  $A_{jk}$  is not a square matrix.) It turns out, at NLO,  $A_{jk}(\mu = m_H) = 0$  [7]; thus  $f_k^{VF}$  are continuous with this choice of matching point. There is a rather significant discontinuity in heavy quark distributions and the gluon distribution at NNLO.

With the matching conditions, Eq. 7,  $\{f_j^{VF}(\mu)\}$  are uniquely defined for all values of  $\mu$ . We shall omit the superscript VF in the following. Moreover, when there is a need to focus on  $f_j(\mu)$  in the vicinity of  $\mu = m_H$ , where there may be a discontinuity, we use  $f_j^{+/-}(\mu)$  to distinguish the above/below branch of the function. As indicated in Eq. 7,  $f_j^-$  correspond to the  $n_f$ -flavor PDFs, and  $f_j^+$  to the  $(n_f + 1)$ -flavor ones.

### 1.3.3 The Structure of a GM VFNS, Minimal Prescription and Additional Freedom

Physical quantities should be independent of the choice of scheme; hence, in a GM VFNS, we must require the theoretical expressions for the structure functions to be continuous across the matching point  $\mu = Q = m_H$  to each order of perturbative theory:

$$F(x, Q) = C_k^-(m_H/Q) \otimes f_k^-(Q) = C_j^+(m_H/Q) \otimes f_j^+(Q) \quad (8)$$

$$\equiv C_j^+(m_H/Q) \otimes A_{jk}(m_H/Q) \otimes f_k^-(Q). \quad (9)$$

where we have suppressed the structure function label ( $\lambda$ ) on  $F$ 's and  $C$ 's, and used the notation  $C_k^{+/-}$  to denote the Wilson coefficient function  $C_k(m_H/Q)$  above/below the matching point respectively. Hence, the GM VFNS coefficient functions are also, in general, discontinuous, and must satisfy the transformation formula:

$$C_k^-(m_H/Q) = C_j^+(m_H/Q) \otimes A_{jk}(m_H/Q). \quad (10)$$

order-by-order in  $\alpha_s$ . For example, at  $\mathcal{O}(\alpha_s)$ ,  $A_{Hg} = \alpha_s P_{qg}^0 \ln(Q/m_H)$ , this constraint implies,

$$C_{H,g}^{-,1}(m_H/Q) = \alpha_s C_{H,H}^{+,0}(m_H/Q) \otimes P_{qg}^0 \ln(Q/m_H) + C_{H,g}^{+,1}(m_H/Q). \quad (11)$$

where the numeral superscript (0,1) refers to the order of calculation in  $\alpha_s$  (for  $P_{jk}$ , the order is by standard convention one higher than indicated), and the suppressed second parton index on the Wilson coefficients (cf. Eq. 2) has been restored to make the content of this equation explicit. Eq. (11) was implicitly used in defining the original ACOT scheme [9]. The first term on the RHS of Eq. 11, when moved to the LHS, becomes the *subtraction term* of Ref. [9] that serves to define the Wilson coefficient  $C_{H,g}^{+,1}(m_H/Q)$  (hence the scheme) at order  $\alpha_s$ , as well as to eliminate the potentially infra-red unsafe logarithm in the gluon fusion term ( $C_{H,g}^{-,1}(m_H/Q)$ ) at high energies.

The GM VFNS as described above, consisting of the general framework of [6, 7], along with transformation matrices  $\{A_{jk}\}$  calculated to order  $\alpha_s^2$  by [12, 13], is accepted in principle by all recent work on PQCD with mass. Together, they can be regarded as the *minimal GM VFNS*.

The definition in Eq. 10 was applied to find the asymptotic limits ( $Q^2/M_H^2 \rightarrow \infty$ ) of coefficient functions in [12, 13], but it is important to observe that it does not completely define all Wilson coefficients across the matching point, hence, there are additional flexibilities in defining a specific scheme [7, 14, 15, 22]. This is because, as mentioned earlier, the transition matrix  $\{A_{jk}\}$  is not a square matrix—it is  $n_f \times (n_f + 1)$ . It is possible to swap  $\mathcal{O}(m_H/Q)$  terms between Wilson coefficients on the right-hand-side of Eq. 10 (hence redefining the scheme) without violating the general principles of a GM VFNS. For instance, one can swap  $\mathcal{O}(m_H/Q)$  terms between  $C_H^{+,0}(m_H/Q)$  and  $C_g^{+,1}(m_H/Q)$  while keeping intact the relation (11) that guarantees the

continuity of  $F(x, Q)$  according to Eq. 8. This general feature, applies to (10) to all orders. It means, in particular, that there is no need to calculate the coefficient function  $C_H^{+,i}(m_H/Q)$ , for any  $i$  – it can be chosen as a part of the definition of the scheme. Also, it is perfectly possible to define coefficient functions which do not individually satisfy the constraint in Eq. 3, since Eq. 10 guarantees ultimate cancellation of any violations between terms. However, this will not occur perfectly at any finite order so modern definitions do include the constraint explicitly, as outlined in Sec. 1.2.4.

The additional flexibility discussed above has been exploited to simplify the calculation, as well as to achieve some desirable features of the prediction of the theory by different groups. Of particular interest and usefulness is the general observation that, given a GM VFNS calculation of  $\{C_j^+\}$ , one can always switch to a simpler scheme with constant  $\{\tilde{C}_j^+\}$

$$\tilde{C}_H^+(m_H/Q) = C_H^+(0) \quad (12)$$

This is because the shift  $(C_H^+(m_H/Q) - C_H^+(0))$  vanishes in the  $m_H/Q \rightarrow 0$  limit, and can be absorbed into a redefinition of the GM scheme as mentioned above. The detailed proof are given in [7, 22]. By choosing the heavy-quark-initiated contributions to coincide with the ZM formulae, the GM VFNS calculation becomes much simplified: given the better known ZM results, we only need to know the full  $m_H$ -dependent contributions from the light-parton-initiated subprocesses; and these are exactly what is provided by the  $n_f$ -flavor FFNS calculations available in the literature. This scheme is known as the *Simplified ACOT scheme*, or SACOT [7, 22].

Further uses of the freedom to reshuffle  $\mathcal{O}(m_H/Q)$  terms between Wilson coefficients, as well as adding terms of higher order in the matching condition (without upsetting the accuracy at the given order) have been employed extensively by the MRST/MSTW group, as will be discussed in Sec. 1.3.5.

#### 1.3.4 CTEQ Implementation of the GM VFNS

The CTEQ group has always followed the general PQCD framework as formulated in [6, 7]. Up to CTEQ6.1, the default CTEQ PDF sets were obtained using the more familiar ZM Wilson coefficients, because, the vast majority of HEP applications carried out by both theorists and experimentalists use this calculational scheme. For those applications that emphasized heavy quarks, special GM VFNS PDF sets were also provided; these were named as CTEQnHQ, where  $n = 4, 5, 6$ .

The earlier CTEQ PDFs are now superseded by CTEQ6.5 [1] and CTEQ6.6 [3] PDFs; these are based on a new implementation of the general framework described in previous sections, plus using the simplifying SACOT choice of heavy quark Wilson coefficients [9, 23] specified by Eq. 12 above. There are no additional modifications of the formulae of the minimal GM VFNS, as described in previous sections. CTEQ uses the convention of designating tree-level, 1-loop, 2-loop calculations as LO, NLO, and NNLO, for all physical quantities,  $F_\lambda^{\text{tot}}$ ,  $F_\lambda^{\text{H}}$ , ... etc., cf. Sec. 1.2.6.

With these minimal choices, this implementation is extremely simple. Continuity of physical predictions across matching points in the scale variable  $\mu = Q$  is guaranteed by Eqs. 8 and 10; and continuity across physical thresholds in the physical variable  $W$ , for producing heavy

flavor final states, are guaranteed by the use of ACOT- $\chi$  rescaling variables 5, as described in Sec. 1.2.4.

For example, to examine the continuity of physical predictions to NLO in this approach, we have, for the below/above matching point calculations:

$$\begin{aligned} F_2^{-H}(x, Q^2) &= \alpha_s C_{2,Hg}^{-,1} \otimes g^{n_f} \\ F_2^{+H}(x, Q^2) &= \alpha_s C_{2,Hg}^{+,1} \otimes g^{n_f+1} + (C_{2,HH}^{+,0} + \alpha_s C_{2,HH}^{+,1}) \otimes (h + \bar{h}) \end{aligned} \quad (13)$$

where non-essential numerical factors have been absorbed into the convolution  $\otimes$ . The continuity of  $F_2^H(x, Q^2)$  in the scaling variable  $\mu = Q$  is satisfied by construction (Eq. 9) because the relation between the PDFs given by Eq. 7 and that between the Wilson coefficients given by Eq. 8 involve the same transformation matrix  $\{A_{jk}\}$  (calculated in [12, 13, 21]). In fact, to this order,  $A_{Hg} = \alpha_s P_{qg}^0 \ln(Q/m_H)$ , hence

$$\begin{aligned} h(\bar{h}) &= 0 \\ g^{n_f+1} &= g^{n_f} \\ C_{2,Hg}^{+,1} &= C_{2,Hg}^{-,1}, \end{aligned}$$

at the matching point  $\mu = Q = m_H$ . Thus, the two lines in Eq. 13 give the same result, and  $F_2^H(x, Q^2)$  is continuous. The separate issue of continuity of  $F_2^H(x, Q^2)$  in the physical variable  $W$  across the production threshold of  $W = 2m_H$  is satisfied automatically by each individual term (using the ACOT- $\chi$  prescription for the quark terms and straightforward kinematics for the gluon term).

In the CTEQ approach, all processes are treated in a uniform way; there is no need to distinguish between neutral current (NC) and charged current (CC) processes in DIS, (among others, as in MRST/MSTW). All CTEQ global analyses so far are carried out up to NLO. This is quite adequate for current phenomenology, given existing experimental and other theoretical uncertainties. Because NNLO results has been known to show signs of unstable behavior of the perturbative expansion, particularly at small- $x$ , they are being studied along with resummation effects that can stabilize the predictions. This study is still underway.

### 1.3.5 MRST/MSTW Implementation of the GM VFNS

**Prescription.** In the Thorne-Roberts (TR) heavy flavour prescriptions, described in [14, 15] the ambiguity in the definition of  $C_{2,HH}^{VF,0}(Q^2/m_H^2)$  was exploited by applying the constraint that  $(dF_2^H/d \ln Q^2)$  was continuous at the transition point (in the gluon sector). However, this becomes technically difficult at higher orders. Hence, in [20] the choice of heavy-flavour coefficient functions for  $F_2^H$  was altered to be the same as the SACOT( $\chi$ ) scheme described above. This choice of heavy-flavour coefficient functions has been used in the most recent MRST/MSTW analysis, in the first instance in [2]. To be precise the choice is

$$C_{2,HH}^{VF,n}(Q^2/m_H^2, z) = C_{2,HH}^{ZM,n}(z/x_{max}). \quad (14)$$

This is applied up to NNLO in [20] and in subsequent analyses. For the first time at this order satisfying the requirements in Eq.(10) leads to discontinuities in coefficient functions, which up

to NNLO cancel those in the parton distributions. This particular choice of coefficient functions removes one of the sources of ambiguity in defining a GM VFNS. However, there are additional ambiguities in the MRST/MSTW convention for counting LO, NLO, ... calculations (cf. Sec.1.2.6), coming about because the ordering in  $\alpha_S$  for  $F_2^H(x, Q^2)$  is different above and below matching points in Eqs. 9-11. (These complications do not arise in the minimal GM VFNS adopted by CTEQ, as already mentioned in the previous subsection.)

For the neutral current DIS  $F_2$  structure function, the above-mentioned ambiguities can be seen as follows:

	below	above
LO	$\frac{\alpha_S}{4\pi} C_{2,Hg}^{-,1} \otimes g^{n_f}$	$C_{2,HH}^{+,0} \otimes (h + \bar{h})$
NLO	$\left(\frac{\alpha_S}{4\pi}\right)^2 (C_{2,Hg}^{-,2} \otimes g^{n_f} + C_{2,Hq}^{-,2} \otimes \Sigma^{n_f})$	$\frac{\alpha_S}{4\pi} (C_{2,HH}^{+,1} \otimes (h + \bar{h}) + C_{2,Hg}^{+,1} \otimes g^{n_f+1})$
NNLO	$\left(\frac{\alpha_S}{4\pi}\right)^3$	$\sum_i C_{2,Hi}^{+,2} \otimes f_i^{n_f} \left(\frac{\alpha_S}{4\pi}\right)^2 \sum_j C_{2,Hj}^{+,2} \otimes f_j^{n_f+1}$ ,

(15)

with obvious generalization to even higher orders. This means that switching directly from a fixed order with  $n_f$  active quarks to fixed order with  $n_f + 1$  active quarks leads to a discontinuity in  $F_2^H(x, Q^2)$ . As with the discontinuities in the ZM-VFNS already discussed this is not just a problem in principle – the discontinuity is comparable to the errors on data, particularly at small  $x$ . The TR scheme, defined in [14, 15], and all subsequent variations, try to maintain the particular ordering in each region as closely as possible. For example at LO the definition is

$$\begin{aligned}
 F_2^H(x, Q^2) &= \frac{\alpha_S(Q^2)}{4\pi} C_{2,Hg}^{-,1}(Q^2/m_H^2) \otimes g^{n_f}(Q^2) \\
 &\rightarrow \frac{\alpha_S(m_H^2)}{4\pi} C_{2,Hg}^{-,1}(1) \otimes g^{n_f}(m_H^2) + C_{2,HH}^{+,0}(Q^2/m_H^2) \otimes (h + \bar{h})(Q^2). \quad (16)
 \end{aligned}$$

The  $\mathcal{O}(\alpha_S)$  term is frozen when going upwards through  $Q^2 = m_H^2$ . This generalizes to higher orders by freezing the term with the highest power of  $\alpha_S$  in the definition for  $Q^2 < m_H^2$  when moving upwards above  $m_H^2$ . Hence, the definition of the ordering is consistent within each region, except for the addition of a constant term (which does not affect evolution) above  $Q^2 = m_H^2$  which becomes progressively less important at higher  $Q^2$ , and whose power of  $\alpha_S$  increases as the order of the perturbative expansion increases.

This definition of the ordering means that in order to define a GM VFNS at NNLO [20] one needs to use the  $\mathcal{O}(\alpha_S^3)$  heavy-flavour coefficient functions for  $Q^2 \leq m_H^2$  (and that the contribution will be frozen for  $Q^2 > m_H^2$ ). This would not be needed in a ACOT-type scheme. As mentioned above, these coefficient functions are not yet calculated. However, as explained in [20], one can model this contribution using the known leading threshold logarithms [24] and leading  $\ln(1/x)$  terms derived from the  $k_T$ -dependent impact factors [25]. This results in a significant contribution at small  $Q^2$  and  $x$  with some model dependence. However, variation in the free parameters does not lead to a large change.<sup>8</sup>

<sup>8</sup>It should be stressed that this model is only valid for the region  $Q^2 \leq m_H^2$ , and would not be useful for a NNLO

The above discussions focused on  $F_2^H$ ; but they mostly apply to  $F_L$  as well. We only need to mention that, with the adoption of the SACOT prescription for heavy-quark initiated contributions (i.e. using the ZM version of the Wilson coefficient),  $F_L^H$  vanishes at order  $\alpha_s^0$  as it does in the TR prescriptions. (This zeroth order coefficient function does appear in some older GM VFNS definitions.) According to the MRST/MSTW convention, the order  $\alpha_s^1$  term of  $F_L$  (both light and heavy flavour) counts as LO, and so on, whereas in the CTEQ convention each relative order is a power of  $\alpha_S$  lower.

The general procedure for the GM VFNS for charged-current deep inelastic scattering can work on the same principles as for neutral currents, but one can produce a *single* charm quark from a strange quark so  $\chi = x(1 + m_c^2/Q^2)$ . However, there is a complication compared to the neutral current case because the massive FFNS coefficient functions are not known at  $\mathcal{O}(\alpha_S^2)$  (only asymptotic limits [27] have been calculated). These coefficient functions are needed in a TR-type scheme at low  $Q^2$  at NLO, and for any GM VFNS at all  $Q^2$  at NNLO. This implies that we can only define the TR scheme to LO and the ACOT scheme to NLO. However, known information can be used to model the higher order coefficient functions similarly to the TR scheme definition to NNLO for neutral currents. A full explanation of the subtleties can be found in [28].

**Scheme variations.** The inclusion of the complete GM VFNS in a global fit at NNLO first appeared in [2], and led to some important changes compared a previous NNLO analysis, which had a much more approximate inclusion of heavy flavours (which was explained clearly in the Appendix of [29]). There is a general result that  $F_2^c(x, Q^2)$  is flatter in  $Q^2$  at NNLO than at NLO, as shown in Fig. 4 of [2], and also flatter than in earlier (approximate) NNLO analyses. This had an important effect on the gluon distribution. As seen in Fig. 5 of [2], it led to a larger gluon for  $x \sim 0.0001 - 0.01$ , as well as a larger value of  $\alpha_S(M_Z^2)$ , both compensating for the naturally flatter evolution, and consequently leading to more evolution of the light quark sea. Both the gluon and the light quark sea were 6 – 7% greater than in the MRST2004 set [30] for  $Q^2 = 10,000 \text{ GeV}^2$ , the increase maximising at  $x = 0.0001 - 0.001$ . As a result there was a 6% increase in the predictions for  $\sigma_W$  and  $\sigma_Z$  at the LHC. This would hold for all LHC processes sensitive to PDFs in this  $x$  range, but would be rather less for processes such as  $t\bar{t}$  pair production sensitive to  $x \geq 0.01$ . This surprisingly large change is a correction rather than a reflection of the uncertainty due to the freedom in choosing heavy flavour schemes and demonstrates that the MRST2004 NNLO distributions should now be considered to be obsolete.

To accompany the MRST 2006 NNLO parton update there is an unofficial “MRST2006 NLO” set, which is fit to exactly the same data as the MRST2006 NNLO set. By comparing to the 2004 MRST set one can check the effect on the distributions due to the change in the prescription for the GM VFNS at NLO without complicating the issue by also changing many other things in the analysis. The comparison of the up quark and gluon distributions for the “MRST2006 NLO” set and the MRST2004 NLO set, i.e. the comparable plot to Fig. 5 of [2] for NNLO, is shown in Fig. 1. As can be seen it leads to the same trend for the partons as at NNLO, i.e. an increase in the small- $x$  gluon and light quarks, but the effect is much smaller –

---

FFNS at all  $Q^2$  since it contains no information on the large  $Q^2/m_H^2$  limits of the coefficient functions. A more general approximation to the  $\mathcal{O}(\alpha_S^3)$  coefficient functions could be attempted, but full details would require first the calculation of the  $\mathcal{O}(\alpha_S^3)$  matrix element  $A_{Hg}$ . This more tractable project is being investigated at present [26].

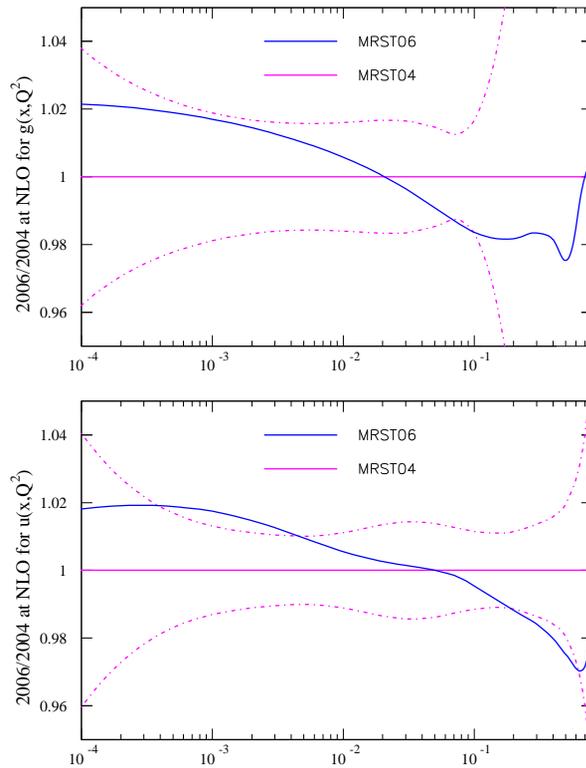


Fig. 1: A comparison of the unpublished “MRST2006 NLO” parton distributions to the MRST2004 NLO distributions. In order to illustrate the significance of the size of the differences, the uncertainty on the MRST2001 distributions is used for the 2004 distributions.

a maximum of a 2% change. Also, the value of the coupling constant increases by 0.001 from the 2004 value of  $\alpha_S(M_Z^2) = 0.120$ . From momentum conservation there must be a fixed point and this is at  $x \sim 0.05$ . Hence,  $W, Z$  and lighter particle production could be affected by up to 2 – 3%, and very high mass states by a similar amount, but final states similar in invariant mass to  $t\bar{t}$  will be largely unaffected. Hence, we can conclude that the change in our choice of the heavy-flavour coefficient function alone leads to changes in the distributions of up to 2%, and since the change is simply a freedom we have in making a definition, this is a theoretical uncertainty on the partons, much like the frequently invoked scale uncertainty. Like the latter, it should decrease as we go to higher orders.

### 1.3.6 Comparisons

We have tried to make clear that both the CTEQ and the MRST/MSTW approaches are consistent with the PQCD formalism with non-zero heavy quark masses  $\{m_H\}$ . In this sense, they are both “valid”. In addition, they both adopt certain sensible practises, such as the numerically significant rescaling-variable approach to correctly treat final-state kinematics (ACOT- $\chi$ ), and the calculationally simplifying SACOT prescription for the quark-parton initiated subprocesses. These common features ensure broad agreement in their predictions. This is borne out by the fact that global QCD analyses carried out by both groups show very good agreement with all available hard scattering data, including the high-precision DIS total inclusive cross sections and semi-inclusive heavy flavor production cross sections; and that the predictions for higher energy cross sections at LHC for the important W/Z production process agree rather well in the most recent versions of these analyses [2, 3].<sup>9</sup> Comparisons of experiment for the abundant data on total inclusive cross sections (and the associated structure functions) with theory are well documented in the CTEQ and MRST/MSTW papers. Here we only show the comparison of the recent H1 data sets on cross sections for charm and bottom production [31] to the latest CTEQ and MSTW calculations. This figure illustrates the general close agreement between the two calculations. (Also, see below.)

Because the main source of the differences between the two implementations arise from the different conventions adopted for organizing the perturbative calculation, it is impossible to make a direct (or clear-cut) comparison between the two calculations. By staying with the conventional order-by-order formulation, the CTEQ approach has all the simplicities of the minimal GM VFNS. With the alternative LO/NLO/NNLO organization, the MRST/MSTW approach includes specifically chosen higher-order terms at each stage of the calculation for different physical quantities (e.g.  $F_2^{\text{tot}}, F_L^{\text{tot}}, F_2^{\text{H}}$ , in Secs. 1.3.5) with their associated Wilson coefficients (e.g. Eqs. 15,16). The choices are a matter of taste because, with the same Wilson coefficients (with heavy quark mass) available in the literature (such as [12, 13]), both analyses can be extended to the appropriate order, and they should contain the same information. So far, MRST/MSTW has carried out their analyses to one order higher than CTEQ. In practice, we have seen one comparison of the “NLO” predictions of the two approaches in Fig. 2 that shows remarkable general agreement with each other, and with experimental data. Some expected differences at small- $x$ , due to the higher order term included in the MRST/MSTW calculation are

<sup>9</sup>Some apparent worrying discrepancies in the predictions for the W/Z cross-sections at LHC between [1] and [30] have been superseded by the recent analyses.

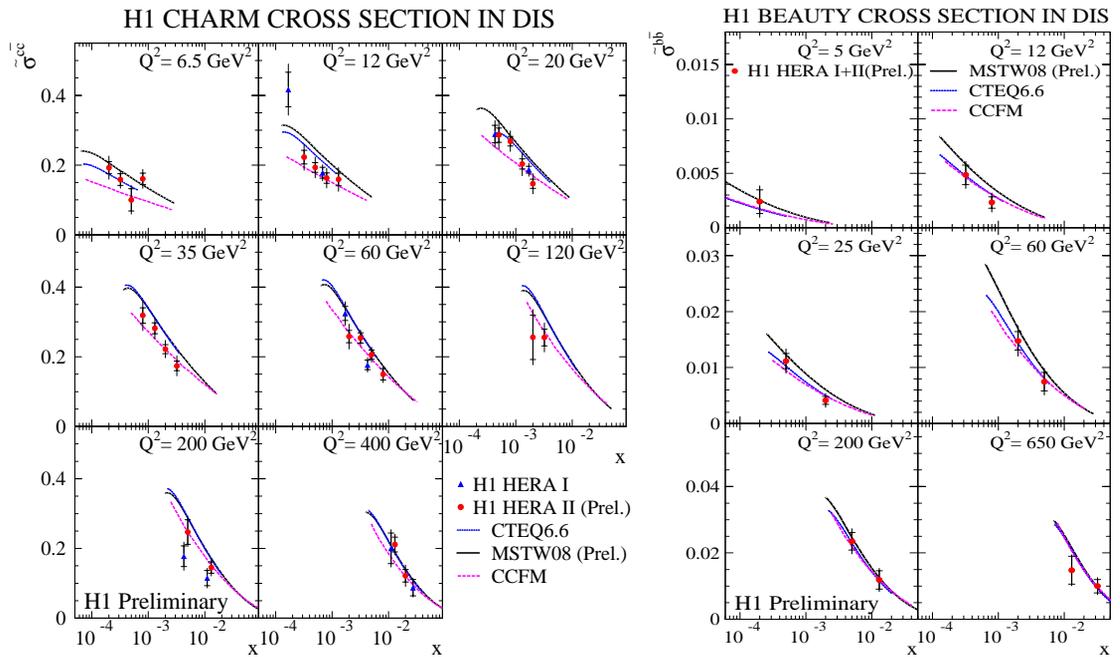


Fig. 2: Comparison of the predictions for  $\bar{\sigma}^{c\bar{c}}(x, Q^2)$  and  $\bar{\sigma}^{b\bar{b}}(x, Q^2)$  compared to preliminary data from H1.

present. Compared to experimental data, the CTEQ curves seem to give a slightly better description of data in this region of difference; but this should not be taken seriously in view of the above discussions. We intend to make a more quantitative study of the differences between the alternative formulations of a GM VFNS and ZM VFNS in a future publication.

#### 1.4 Use of Parton Distribution Functions

Some commonly asked questions in the user community for PDFs are along the lines of: (i) Which available PDF set is most appropriate for my particular calculation? and (ii) If PDF set A was obtained using scheme A (say, ZMVFNS/GMVFNS-MSTW/GMVFNS-CTEQ) do I have to use the same scheme A for my Wilson coefficients (otherwise my calculation would be inconsistent)? Whereas it is impossible to answer all such questions at once, the following observations should provide useful guidelines toward the appropriate answers. Foremost, it is important to bear in mind that in the perturbative approach, all calculations are approximate; hence the goodness of the approximation is the most (or only) relevant consideration. Any fast, or absolute, rules or prescriptions would be misguided.

\* For applications at very high energy scales, e.g. most LHC processes, it is perfectly fine to use the ZM formulae for the hard-scattering coefficient *irrespective of the choice of PDF sets* (see below), since the ZM Wilson coefficients are good approximations to the GM ones (valid to  $\mathcal{O}(M^2/Q^2)$  where  $M$  represents the typical mass in the relevant parton subprocess—heavy quarks or other produced particles), and the ZM coefficients are much simpler and *much more* readily available.

On the other hand, for applications involving physical scales  $Q \sim \mathcal{O}(M)$ , such as comparison to precision DIS data at HERA, it is important both to use GM Wilson coefficients, and to ensure that these are consistent with those adopted in generating the PDF set to be used in the calculation.

\* For the global analyses that yield the PDF sets, it matters whether the ZM VFNS or GM VFNS scheme is used in the calculation, since a substantial fraction of the input DIS data are in the region where  $Q$  is not very large compared to the heavy quark masses  $m_{c,b}$  (the top quark does not play a significant role in these analysis). Thus, the ZM-VFNS and GM VFNS PDFs can differ in some  $x$ -range, even if they agree quite well in general (cf. [1]). For example, the widely used CTEQ6.1 (ZM-VFNS) and the most recent CTEQ6.5/CTEQ6.6 (GM VFNS) PDF sets both give excellent fits to the available data, yet the differences (mainly around  $x \sim 10^{-3}$ ) are enough to lead to a 6% shift in the predictions for cross sections for  $W, Z$  and similar mass states at the LHC. Higher mass final states are much less affected.

The above differences arise from two sources: (i) the treatment of final-state counting (Sec. 1.2.3) and phase space (Sec. 1.2.4); and (ii) mass effects in the Wilson coefficients. The first is numerically significant for reasons explained in those sections, and it can potentially be removed to produce an improved ZM VFNS (Sec. 1.3.1).

\* The differences between PDFs obtained using different GM VFNS implementations, such as those by CTEQ and MSTW groups discussed in the main part of this review, are much smaller than those between the ZM and GM VFNS. This is because the treatments of final states are similar, and the differences in the Wilson coefficients are much reduced also. The current NLO predictions on  $W/Z$  cross sections at LHC by the CTEQ and MSTW groups, for instance, are within 2% [4].

\* What about single-flavor (say,  $n_f$ ) FFNS PDFs that are commonly believed to be needed for FFNS calculations, such as for heavy flavor production processes? We would like to point out, perhaps surprisingly to many readers, that: (i) with the advent of GM VFNS PDFs, *the FFNS PDFs are not in principle needed* for consistency; and (ii) the use of  $n_f$ -flavor FFNS PDFs in a  $n_f$ -flavor calculation is much *less reliable* than using the GM VFNS (if the latter is available). The reasons for these assertions are fairly easy to see, as we now explain.

First of all, as we emphasized in Sec. 1.2.2, the GM VFNS is, by definition, a composite scheme that *is* the  $n_f$ -FFNS within the region of validity of the latter. In principle one *can* use the GM VFNS PDFs in the FFNS calculations within the region where the FFNS is reliable. (In practice this range of validity (in energy scale  $\mu$ ) extends up to several times  $m_H$ , cf. second to last paragraph of Sec. 1.2.2.) Secondly, since any given  $n_f$ -FFNS has only a limited range of validity (Sec. 1.2.2), the global analysis used to determine any  $n_f$ -FFNS PDF set is inherently a compromise. This compromise is likely to be a fairly bad one for two reasons. Firstly, the limited range of validity implies that only a fraction of the data used in the global analysis can be legitimately applied. If one excludes all the data outside of the region of validity of the theory (not an easily-defined region), the constraining power of the analysis would greatly suffer. If, instead, one includes all the points in the analysis anyway, the PDFs will compensate, much like the case of the fit using the basic ZM VFNS. This can result in a good comparison to data (as in the ZM VFNS [32]), but this is potentially misleading since the compensation is caused by

the wrong physics. In either of the cases, the PDFs resulting from a fit using the FFNS will be unreliable. Secondly, Wilson coefficients in the FFNS only exist for the DIS process beyond LO, hence the ZM approximation to  $n_f$ -FFNS must be used. We note, although this second point is shared by current GM VFNS analyses, the ZM VFNS approximation to GM VFNS is a much better approximation than that of ZM FFNS to  $n_f$ -FFNS. (For instance, for collider jet data sets, the ZM 3- or 4-flavor calculation would be way-off the correct one. This is not a problem for the GM VFNS case.) These inherent problems motivated an alternative approach to FFNS PDFs in [33]: rather than performing a (imperfect) FFNS global fit, one simply generates them by fixed  $n_f$ -flavor QCD evolution from a set of initial PDFs obtained in an existing (bona fide) GM VFNS global analysis! Because of the different QCD evolution, however, the PDFs will be different from the original GM VFNS ones crossing heavy flavor thresholds; and the fits to the global data will correspondingly deteriorate, particularly for the high precision HERA data sets at higher  $Q^2$ . Thus, these PDFs deviate from truth in a different way. The relative merit between this approach and the conventional FFNS global fits is difficult to gauge because there are no objective criteria for making the assessment.

Returning to the original question that started this bullet item, we can summarize the options available to match PDFs with a FFNS calculation such as HQVDIS [34] for heavy quark production: (i) conventional FFNS PDFs (CTEQ, GRV), suitably updated if necessary [35]; (ii) PDFs generated by FFNS evolution from GM VFNS PDFs at some initial scale  $Q_0$  (MSTW [33], but also can easily be done with CTEQ); or, (iii) simply use the most up-to-date GM VFNS PDFs (MSTW, CTEQ) for all  $Q$ . For reasons discussed in the previous paragraphs, each option has its advantages and disadvantages. (i) and even (ii) are theoretically self-consistent, while (iii) is not, e.g. it opens up the awkward question of how many flavours to use in the definition of  $\alpha_s$ . However, the PDFs in (iii) are intrinsically much more accurately and precisely determined. Hence, in practical terms it is not obvious which would be most “correct”.<sup>10</sup> The choice reduces to a matter of taste, and for some, of conviction. The differences in results, obtained using these options, should not be too large, since they are mostly of one order higher in  $\alpha_s$ ; and, in an approximate manner, they define the existing theoretical uncertainty. In principle, an approach that combines the advantages of all three, hence could work the best, would be to use PDFs obtained in the GM VFNS, but with the transition scale  $\mu_T$  (Sec. 1.2.3) set at a much higher value than  $m_H$  for each heavy flavor threshold. But this option is rather cumbersome to implement (as has been hinted in Sec. 1.2.3), hence has not been done.

\* There exists another class of applications, involving multiple-scale processes, such as heavy flavor production at hadron colliders with finite transverse momentum  $p_T$  or in association with  $W/Z$  or Higgs, for which PQCD calculations are more complex than the familiar one-hard-scale case, as implicitly assumed above. Since these processes can play an important role in LHC, there has been much discussions, and controversies, in recent literature about the various approaches that may be applied [36]. Both the GM VFNS [37] and FFNS approaches have been advocated [38]. The problem is complex, generally because more than one kind of potentially large logarithms occur in these problems, and they cannot be effectively controlled all at once with some suitable choice of scheme. A detailed discussion is outside the scope of this paper,

---

<sup>10</sup>Although it is certainly better to use a current GM VFNS set of PDFs than an out-of-date FFNS set.

although our remark about the FFNS PDFs above could be helpful (and relieve some of the anxieties expressed in the literature).

All in all, for general applications, taking into account all the considerations above, the modern GM VFNS PDF sets are clearly the PDFs of choice.

## 1.5 Intrinsic Heavy Flavour

Throughout the above discussions we have made the assumption that all heavy quark flavour is generated from the gluon and lighter flavours through the perturbative QCD evolution, starting from the respective scale  $\mu = m_H$ . This is usually referred to as the *radiatively generated heavy flavor* scenario. From the theoretical point of view, this is reasonable for heavy flavors with mass scale ( $m_H$ ) very much higher than the on-set of the perturbative regime, say  $\sim 1$  GeV. Thus, while this assumption is usually not questioned for bottom and top, the case for charm is less obvious. In fact, the possibility for a non-negligible *intrinsic charm* (IC) component of the nucleon at  $\mu = Q \sim m_c$  was raised a long time ago [39]; and interests in this possibility have persisted over the years. Whereas the dynamical origin of such a component can be the subject of much debate, the phenomenological question of its existence can be answered by global QCD analysis: do current data support the IC idea, and if so, what is its size and shape? This problem has been studied recently by a CTEQ group [40], under two possible scenarios: IC is enhanced at high values of  $x$  (suggested by dynamical models such as [39]), or it is similar in shape to the light-flavor sea quarks (similar to, say, strange). They found that current data do not tightly constrain the charm distribution, but they *can* place meaningful bounds on its size. Thus, while the conventional radiatively generated charm is consistent with data, IC is allowed in both scenarios. For the model-inspired (large- $x$ ) case, the size of IC can be as large as  $\sim 3$  times that of the crude model estimates, though comparison to the EMC  $F_2^c$  data [41] imply contributions somewhat smaller [42]. If such an IC component does exist, it would have significant impact on LHC phenomenology for certain beyond SM processes. For the sea-like IC case, the bound on its size is looser (because it can be easily interchanged with the other sea quarks in the global fits); its phenomenological consequences are likewise harder to pin-point.

From a theoretical point of view, intrinsic heavy flavour and GM VFNS definitions were discussed in [43]. Allowing an intrinsic heavy quark distribution actually removes the redundancy in the definition of the coefficient functions in the GM VFNS, and two different definitions of a GM VFNS will no longer be identical if formally summed to all orders, though they will only differ by contributions depending on the intrinsic flavour. Consider using identical parton distributions, including the intrinsic heavy quarks, in two different flavour schemes. The heavy-quark coefficient functions at each order are different by  $\mathcal{O}(m_H^2/Q^2)$ . This difference has been constructed to disappear at all orders when combining the parton distributions other than the intrinsic heavy quarks, but will persist for the intrinsic contribution. The intrinsic heavy-flavour distributions are of  $\mathcal{O}(\Lambda_{QCD}^2/m_H^2)$ , and when combined with the difference in coefficient functions the mass-dependence cancels leading to a difference in structure functions of  $\mathcal{O}(\Lambda_{QCD}^2/Q^2)$ . It has been shown [7] that for a given GM VFNS the calculation of the structure functions is limited in accuracy to  $\mathcal{O}(\Lambda_{QCD}^2/Q^2)$ . Hence, when including intrinsic charm, the scheme ambiguity is of the same order as the best possible accuracy one can obtain in leading twist QCD, which is admittedly better than that obtained from ignoring the intrinsic heavy flavour (if it exists) as  $Q^2$

increases above  $m_H^2$ . It is intuitively obvious that best accuracy will be obtained from a definition of a GM VFNS where all coefficient functions respect particle kinematics. In fact, the most recent CTEQ and MSTW prescriptions would provide identical contributions to the structure functions from the same intrinsic charm parton distribution.

**Acknowledgements** We thank Matteo Cacciari for his unrelenting efforts to bring about this review on behalf of the Heavy Flavor Workshop of the HERALHC Workshop. We thank our collaborators in CTEQ and MRST/MSTW for collaborations which underlies much of the content of this paper. WKT would like especially to acknowledge the insight provided by John Collins on the theoretical foundation of PQCD with heavy quarks. We also thank Pavel Nadolsky, Fred Olness, Ingo Schienbein, Jack Smith and Paul Thompson for reading the manuscript and making useful comments that brought about improvements in the presentation of the paper.

The work of WKT is supported by the National Science Foundation (USA) under the grant PHY-0354838.

## 2 Charmed-meson fragmentation functions with finite-mass corrections

*Authors: B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger*

A straight-forward and conventional approach to include heavy-quark mass effects in the theoretical predictions for the production of single heavy-flavor mesons consists in taking into account the non-zero quark mass  $m_h$  in a calculation where only light quarks and the gluon exist in the initial state and the heavy quark is pair-produced in the hard scattering process. Such a scheme is called a fixed-flavor-number scheme (FFNS) and can be implemented, presently, only at NLO. It is reliable in a kinematic region not far above production threshold. At high scales  $\mu$ , however, the presence of logarithmic terms proportional to  $\log(\mu/m_h)$  makes the predictions of a calculation in the FFNS unreliable. These logarithmic terms have to be resummed, which is conventionally done in the so-called zero-mass variable-flavor-number scheme (ZM-VFNS) where the heavy quark is treated as a parton, in addition to light quarks and the gluon. Heavy quark parton distribution functions and fragmentation functions, which are present in this scheme, can absorb the large logarithmic terms and resummation is performed with the help of the DGLAP evolution equations.

The general-mass variable-flavor-number scheme (GM-VFNS) provides a framework for the theoretical description of the inclusive production of single heavy-flavored hadrons, combining virtues of both the FFNS and the ZM-FVNS in a unified approach. It resums large logarithms by the DGLAP evolution of non-perturbative fragmentation functions, guarantees the universality of the latter as in the ZM-VFNS, and simultaneously retains the mass-dependent terms of the FFNS without additional assumptions. It was elaborated at next-to-leading order (NLO) for photo- [45] and hadroproduction [46,47] and  $e^+e^-$  annihilation [48].

Recent progress in the implementation of the GM-VFNS at NLO allowed us to extract mass-dependent FFs for  $D$ -mesons from global fits to  $e^+e^-$  annihilation data [48]. We used experimental data from the Belle, CLEO, ALEPH, and OPAL Collaborations [44]. The fits for  $D^0$ ,  $D^+$ , and  $D^{*+}$  mesons using the Bowler ansatz [49] yielded  $\chi^2/\text{d.o.f.} = 4.03, 1.99, \text{ and } 6.90$ , respectively. The result of the fit for  $D^+$  mesons is shown in Fig. 3.

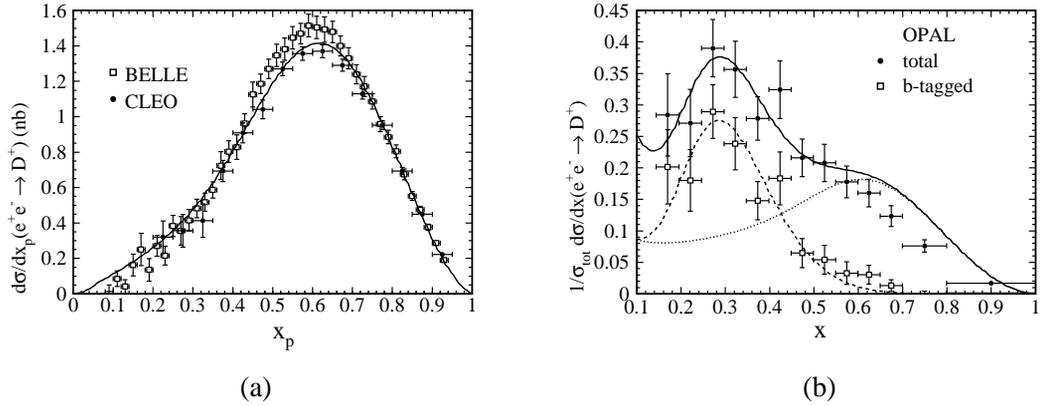


Fig. 3: Comparison of (a) Belle, CLEO, and (b) OPAL data on  $D^+$  mesons [44] with a global fit. The dotted line in panel (b) refers to the  $c$ -quark-initiated contribution.

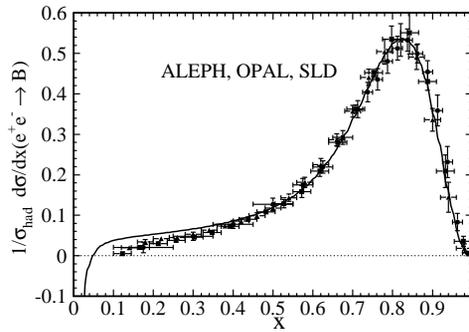


Fig. 4: Comparison of the ALEPH, OPAL, and SLD data on  $B$  meson production [50] with a fit for the  $b \rightarrow B$  FF.

The significance of finite-mass effects can be assessed through a comparison with a similar analysis in the ZM-VFNS. It turned out that for the experimental conditions at Belle and CLEO, charmed-hadron mass effects on the phase space are appreciable, while charm-quark mass effects on the partonic matrix elements are less important. In Figs. 3(a) and (b), the scaled-momentum distributions from Belle and CLEO and the normalized scaled-energy distributions from OPAL for  $D^+$  mesons are compared to the global fits. The Belle and CLEO data prefer higher values for the average  $x$  of the  $c \rightarrow D$  FFs. Due to their smaller errors they dominate the global fit, and the ALEPH and OPAL data are less well described. Charmed hadrons may also originate indirectly through the fragmentation of a  $b$  quark. Our ansatz includes non-perturbative  $b \rightarrow D$  FFs, but these are only weakly constrained by the Belle and CLEO data.

Previous fits of the  $b \rightarrow B$  FFs in the ZM-VFNS [52] were based on  $e^+e^-$  data from ALEPH, OPAL and SLD [50] and used the Kartvelishvili-Likhoded ansatz [53]. As a recent improvement we adjusted the value of  $m_b$  and the energy scale where the DGLAP evolution starts,

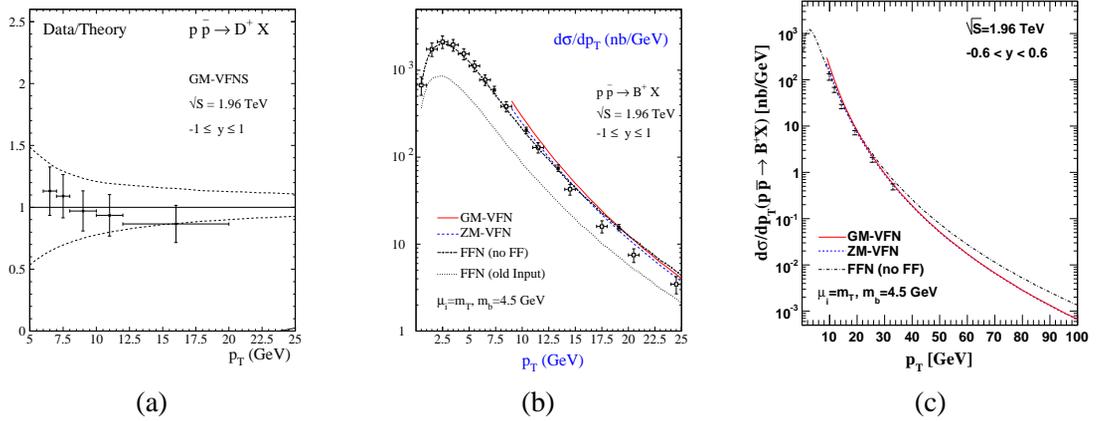


Fig. 5: (a) Comparison of CDF II data [51] on  $D^+$  mesons with the global fit.

to conform with modern PDF sets. The data are well described by the fit, with a  $\chi^2/\text{d.o.f.} = 1.495$ . The result is shown in Fig. 4.

Usage of these new FFs leads to an improved description of the CDF data for charmed-meson production [54] from run II at the Tevatron, as may be seen by comparing Fig. 5(a) in this chapter with Fig. 2(b) of Ref. [47]. Also predictions for  $B$ -meson production agree with CDF II data [51]. Comparing massless and massive calculations, we found that finite- $m_b$  effects moderately enhance the  $p_T$  distribution; the enhancement amounts to about 20% at  $p_T = 2m_b$  and rapidly decreases with increasing values of  $p_T$ , falling below 10% at  $p_T = 4m_b$  (see Fig. 5b). Such effects are thus comparable in size to the theoretical uncertainty due to the freedom of choice in the setting of the renormalization and factorization scales. At higher values of the transverse momentum,  $p_T$ , the predictions of the GM-VFNS and ZM-VFNS approach each other by construction. There, resummation of large logarithms will be important and a FFNS calculation will become inappropriate since it does not resum large logarithms. CDF data reach up to 40 GeV and preliminary data at the highest values of  $p_T$  indicate that resummation of large logarithmic terms will be necessary to obtain a reasonable description of experimental results.

### 3 Fragmentation of heavy quarks with an effective strong coupling constant

*Authors: G. Corcella and G. Ferrera*

We describe a model to include non-perturbative corrections to heavy-quark fragmentation, based on next-to-next-to-leading logarithmic threshold resummation and an effective QCD coupling constant not containing the Landau pole. Comparison with experimental data is also presented.

The hadronization of partons into hadrons cannot be calculated from first principles, but it is usually described in terms of phenomenological models, containing few parameters which need to be tuned to experimental data. In this paper we propose a different approach to describe heavy-quark (bottom and charm) fragmentation in  $e^+e^-$  processes: we use a non-perturbative model [55, 56] including power corrections via an effective strong coupling constant, which does

not exhibit the Landau pole. The interesting feature of such a model is that it does not contain any extra free parameter to be fitted to the data, besides the ones entering in the parton-level calculation. In [57, 58] such a model was also employed in the framework of  $B$ -meson decays and it was found good agreement with the experimental data. Moreover, it was even possible to extract  $\alpha_S(m_Z)$  and the Cabibbo-Kobayashi-Maskawa matrix element  $|V_{ub}|$  from such data [57, 58]. In the following, we shall consider heavy-quark production in  $e^+e^-$  annihilation, in particular  $b$ - and  $c$ -quark production at LEP. In [56], charm-quark fragmentation at the  $\Upsilon(4S)$  resonance was also investigated.

The perturbative fragmentation approach [59], up to power corrections, factorizes the energy distribution of a heavy quark as the convolution of a process-dependent coefficient function, associated with the emission off a massless parton, and a process-independent perturbative fragmentation function, expressing the transition of the light parton into a heavy quark. The heavy-quark spectrum reads:

$$\frac{1}{\sigma} \frac{d\sigma}{dx}(x, Q, m_q) = C(x, Q, \mu_F) \otimes D(x, \mu_F, m_q) + \mathcal{O}((m_q/Q)). \quad (17)$$

where  $Q$  is the hard scale of the process,  $x$  is the heavy-quark energy fraction in the centre-of-mass frame, i.e.  $x = \frac{2E_q}{Q}$ , and  $\mu_F \sim Q$  is the factorization scale.

The perturbative fragmentation function follows the DGLAP evolution equations. As in [55, 56], we use coefficient function and initial condition at next-to-leading order (NLO) and solve the DGLAP equations with a NLO kernel<sup>11</sup>. This way, one resum the large mass logarithms  $\sim \ln(Q^2/m_q^2)$  in the next-to-next-to-leading logarithmic approximation [59]. Furthermore, both coefficient function and initial condition contain terms,  $\sim 1/(1-x)_+$  and  $\sim [\ln(1-x)/(1-x)]_+$ , enhanced when  $x$  approaches 1, which corresponds to soft- or collinear-gluon radiation. One needs to resum such contributions to all orders to improve the perturbative prediction (threshold resummation) [60]. In our analysis, we implement threshold resummation in the next-to-next-to-leading logarithmic (NNLL) approximation. following the general method of [61, 62].

Let us now briefly discuss the phenomenological model which includes non-perturbative power corrections through an effective QCD coupling [55–57, 63]. We start by constructing a general analytic QCD coupling  $\tilde{\alpha}_S(Q^2)$  from the standard one, by means of an analyticity requirement:  $\tilde{\alpha}_S(Q^2)$  is defined to have the same discontinuity as the standard coupling and no other singularity [64]. The coupling constant constructed in this way exhibits no Landau pole, which is subtracted by a power correction, while it has the same discontinuity as the standard one for  $Q^2 < 0$ , related to gluon branching. As discussed in [55], since heavy quark fragmentation is a time-like process, we have to include the absorptive parts of the gluon polarization function into the effective coupling: that amounts to a resummation of constant terms to all orders. As detailed in [55, 56], the effective time-like coupling  $\tilde{\alpha}_S(Q^2)$  is thus given by an integral over the discontinuity of the gluon propagator, with the analytic coupling  $\tilde{\alpha}_S(Q^2)$  entering in the integrand function. At one-loop, for example, one obtains the following effective time-like coupling constant:

$$\tilde{\alpha}_S(Q^2) = \frac{1}{\beta_0} \left[ \frac{1}{2} - \frac{1}{\pi} \arctan \left( \frac{\log(Q^2/\Lambda^2)}{\pi} \right) \right]. \quad (18)$$

---

<sup>11</sup>One could go beyond such a level of accuracy and include next-to-next-to-leading order (NNLO) corrections to the coefficient function, initial condition and to the non-singlet splitting functions.

Our model simply replaces the standard  $\alpha_S(Q^2)$  with the effective time-like coupling constant. As in [55, 56],  $\tilde{\alpha}_S(Q^2)$  is evaluated up to NNLO, i.e. three-loop accuracy. We stress that, even if our model does not contain any free parameter to be fitted to data, we had to choose among possible different prescriptions, mostly concerning the low-energy behaviour of the effective coupling constant. The model presented in [55, 56] is the one which best describes the experimental data.

In Fig. 6 we compare the predictions of the effective-coupling model with experimental data from ALEPH [65], OPAL [66] and SLD [67] on  $B$ -hadron production at the  $Z^0$  pole, and from ALEPH on  $D^{*+}$  production [68]. We learn from the comparison that our model, without introducing any tunable parameter, manages to give a good description of the experimental data. As discussed in [55, 56], even the moments of the  $B$ - and  $D$ -hadron cross section are reproduced quite well.

In summary, we managed to construct a simple non-perturbative model which is able to describe data from rather different processes, namely  $B$ -decays and bottom/charm fragmentation, involving pretty different hard scales. We believe that such results are highly non trivial and that our model deserves further extension to hadron-collider physics. This is in progress.

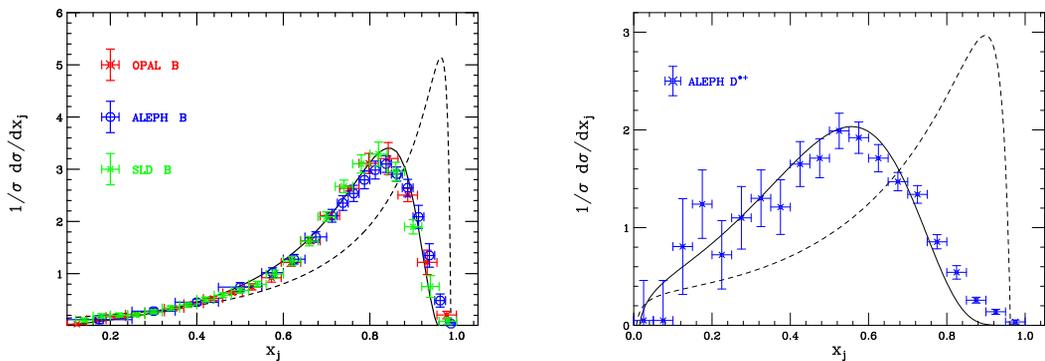


Fig. 6: Results on bottomed (left) and charmed (right) hadron production (solid line), according to the effective-coupling model, compared with the pure parton-level calculation (dashes) and with experimental data.  $x_j$  is the hadron ( $j = B, D$ ) or quark ( $j = b, c$ ) energy fraction at the  $Z^0$  pole.

## 4 Infrared safe determination of jet flavour: theory and applications

Author: A. Banfi and G. Zanderighi

### 4.1 Problems in defining the flavour of a jet

Jets are so far the best-known way to map a complicated event, characterised by a high particle multiplicity, to a simpler one made up of a small number of clusters of particles, *jets*, whose energy-momentum flow is close to that of the original event. By “close to” we mean that jets have to be infrared and collinear (IRC) safe objects, that is their momenta should not change after an extremely soft particle has been added to the event or if any of the particles in the event splits into a quasi-collinear pair. With this requirement jet cross sections can be safely computed in perturbative (PT) QCD. Furthermore, given a partonic event, any IRC safe jet algorithm, in

the soft/collinear limit, does provide a unique mapping to the underlying hard event.<sup>12</sup> It is interesting to investigate whether jet algorithms can be extended so as to define also the flavour of a jet. More precisely, suppose we have a hard event and a new event obtained from the hard event via an arbitrary number of soft emissions and/or collinear splittings. Is it possible to cluster the new event into jets, such that not only the momenta, but also the flavour of the jets, are equal to those of the particles constituting the original hard event?

Attempts to answer this questions have been performed by different experimental groups, whose definitions of jet flavour are based either on the kinematical properties [69] or on the charge of a jet [70]. Although of considerable practical usefulness, these procedures all suffer from IRC unsafety (see [71] for a discussion on this point).

To see where IRC safety problems may arise we need first to introduce our definition of jet flavour. The flavour of a jet is defined as a  $n_f$ -dimensional vector containing in the entry  $f$  the net number of quarks (number of quarks minus number of antiquarks) of flavour  $f$ . A gluon jet will have a flavour vector in which all entries are zero. A clear source of IR unsafety is gluon splitting into a quark and an antiquark that are recombined with different jets, thus changing the underlying jet flavour. At next-to-leading order (NLO), the only singular contribution occurs when the quarks are collinear. In this case, the  $q\bar{q}$  pair is always recombined in the same jet by any IRC safe jet-algorithm, and the resulting jet flavour is also IRC safe. Starting from the next perturbative order however a soft large-angle gluon splitting may produce a  $q$  and  $\bar{q}$  which are both soft but may not be collinear. Therefore the two fermions can be clustered into two different jets, thereby modifying the flavour of those hard jets. In the next section we will analyse specifically the  $k_t$  algorithm, show that its standard version is not IR safe with respect to the jet flavour, and we will see how it can be modified to achieve an IR safe jet-flavour algorithm.

## 4.2 IR safe jet-flavour algorithms at parton level

Let us see how a jet-flavour algorithm should work in the specific case of  $e^+e^-$  annihilation into hadrons. There we consider close-to-Born events with a hard  $q\bar{q}$  pair accompanied by an arbitrary number of soft/collinear partons. One of such configurations is represented in fig. 7. It contains a hard  $q\bar{q}$  pair (at the bottom of each diagram) accompanied by a soft gluon and a soft  $q\bar{q}$  pair originated by the splitting of a large-angle gluon. If one applies the  $k_t$  algorithm [72–74] to such a configuration, to all pairs of particles  $p_i, p_j$  one associates a distance

$$d_{ij} = 2(1 - \cos \theta_{ij}) \times \min\{E_i^2, E_j^2\}, \quad (19)$$

and clusters together the pair whose  $d_{ij}$  is minimum. The resulting set of distances is represented in the picture on the left hand side of fig. 7, where a thick line represents a large distance, while small distances are represented by thin lines. The only large distance obtained with the traditional  $k_t$  algorithm is that between the hard  $q\bar{q}$  pair, while all other distances are small. This is because all other pairs involve at least one soft parton and the distance in eq. (19) depends on the energy of the softest particle only. Looking in particular at the soft  $q$  and  $\bar{q}$ , they can be clustered in different jets thus giving either a couple of gluon jets or two multi-flavoured jets, i.e. not corresponding

<sup>12</sup>Beyond the soft/collinear limit, such a mapping is intrinsically ambiguous due to the presence of interference terms.

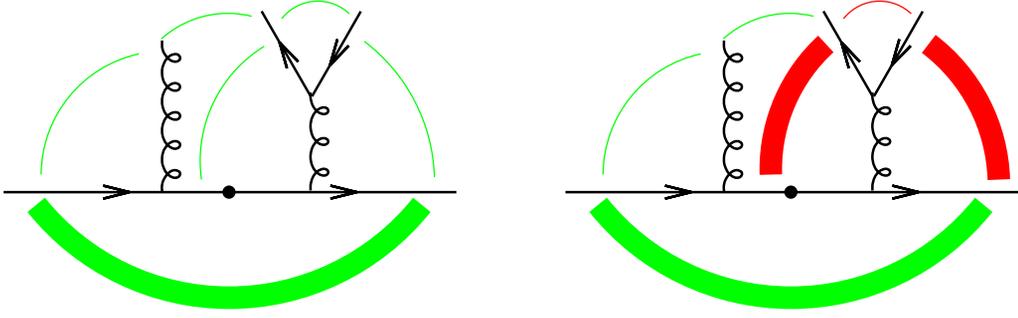


Fig. 7: Pictorial representation of recombination distances for a sample partonic final state in  $e^+e^-$  annihilation in the case of the traditional  $k_t$  algorithm (left) and a  $k_t$ -flavour algorithm (right).

to any QCD parton. The latter case can be eliminated by allowing only recombinations of  $q\bar{q}$  pairs of the same flavour, but the problem of generating fake gluon jets remains. The origin of the problem is that the distance in eq. (19) is modelled so as to compensate the soft and collinear divergence in the matrix element for gluon emission. The  $q\bar{q}$  splitting probability has no soft divergence, so that, without endangering the IRC safety of the algorithm, one could modify the distance in eq. (19) as follows:

$$d_{ij} = 2(1 - \cos \theta_{ij}) \times \begin{cases} \min\{E_i^2, E_j^2\} & \text{softer of } i, j \text{ flavourless,} \\ \max\{E_i^2, E_j^2\} & \text{softer of } i, j \text{ flavoured.} \end{cases} \quad (20)$$

What happens in this case is represented in the picture on the right-hand side of fig. 7, where the new distances are highlighted in red. There, the distance between the soft  $q\bar{q}$  pair is still small, what becomes large is the distance between either of the two and the hard  $q\bar{q}$  pair. In this way soft  $q\bar{q}$  pairs are first recombined together, and only after recombination is the resulting gluon jet recombined with other hard jets. It can be proven that with this modification the resulting flavour determination is IRC safe to all orders in perturbation theory [71].

One can generalise eq. (20) to hadron hadron collisions, defining for each pair of particles a distance parameterised by a jet radius  $R$ :

$$d_{ij} = \frac{\Delta R_{ij}^2}{R^2} \times \begin{cases} \min\{p_{t,i}^2, p_{t,j}^2\} & \text{softer of } i, j \text{ flavourless,} \\ \max\{p_{t,i}^2, p_{t,j}^2\} & \text{softer of } i, j \text{ flavoured,} \end{cases} \quad (21)$$

where  $\Delta R_{ij}^2$  is any collinear safe distance in the rapidity-azimuth  $y$ - $\phi$  plane, for instance  $(y_i - y_j)^2 + (\phi_i - \phi_j)^2$ . Furthermore, to obtain a full flavour determination, one has to add a distance between each particle and the two beams  $B$  and  $\bar{B}$  at positive and negative infinite rapidity respectively. This is achieved by introducing a rapidity dependent transverse momentum for each beam  $p_{t,B}(y)$ ,  $p_{t,\bar{B}}(y)$ , and defining

$$d_{iB} = \begin{cases} \min\{p_{t,i}^2, p_{t,B}^2(y_i)\} & i \text{ flavourless,} \\ \max\{p_{t,i}^2, p_{t,B}^2(y_i)\} & i \text{ flavoured,} \end{cases} \quad (22)$$

and analogously for  $d_{i\bar{B}}$ . The beam hard scales  $p_{t,B}(y)$  and  $p_{t,\bar{B}}(y)$  have to be constructed in such a way that emissions collinear to  $B$  or  $\bar{B}$  get recombined with the right beam, and that

$p_{t,B}(y)$  and  $p_{t,\bar{B}}(y)$  approach the hard scale of the event for central emissions. This is achieved for instance by defining

$$\begin{aligned} p_{t,B}(y) &= \sum_i p_{ti} \left( \Theta(y_i - y) + \Theta(y - y_i) e^{y_i - y} \right) , \\ p_{t,\bar{B}}(y) &= \sum_i p_{ti} \left( \Theta(y_i - y) e^{y - y_i} + \Theta(y - y_i) \right) . \end{aligned} \quad (23)$$

If applied at parton level, these jet-flavour algorithms have two main applications. First of all they can be used in a NLO calculation to assign each event to an underlying Born subprocess. This is needed to correctly merge real and virtual contributions when matching NLO and resummed calculations [75]. A second application of jet-flavour algorithms is the combination of parton showers and matrix elements [76, 77]. For instance, in the CKKW approach [76], the correct Sudakov form factor to be associated to each event is decided only after having clustered the event into jets. This Sudakov form factor depends on the colour charge of the hard emitters, and is therefore correctly computed only if a flavour has been properly (i.e. in a IRC safe way) assigned to each jet.

At hadron level, in general, it is not sensible to distinguish quarks and gluons. However, there is a case in which the flavour algorithm can be successfully applied also at hadron level, that is in the case of heavy flavour production. There all hadrons containing a heavy quark (of the selected flavour) are treated as flavoured, while all other hadrons are considered flavourless. As we will see in the next section, an IRC safe jet-flavour algorithm can thus be exploited to obtain accurate QCD predictions for  $b$ -jet cross sections.

### 4.3 Accurate QCD predictions for $b$ -jet cross sections

A basic measurement in  $b$  production in hadronic collisions is  $b$ -jet transverse momentum spectra. Experimentally a  $b$ -jet is defined as any jet containing at least one  $b$ -flavoured hadron [78]. It is clear that such a definition is collinear unsafe, because any jet containing a  $b\bar{b}$  pair, which should be considered a gluon jet, would be classified as a quark jet. This gives rise to collinear singular contributions if the  $b\bar{b}$  pair arise from a gluon collinear splitting. The resulting collinear singularity is regularised by the  $b$ -quark mass, giving rise to large logarithms at most of relative order  $\alpha_s^n \ln^{2n-1}(p_t/m_b)$ . These gluon splitting (GSP) processes constitute the dominant source of  $b$ -jets at the Tevatron. This is awkward since jets from GSP do not even correspond to one's physical idea of a  $b$ -jet. There are two other production channels, flavour excitation (FEX) and flavour creation (FCR). In FEX one of the constituents of a produced  $b\bar{b}$  pair is collinear to the beam, while the other builds up the  $b$ -jet. This process also contains collinear singularities, which at all orders give rise to terms  $\alpha_s^n \ln^n(p_t/m_b)$ . FCR is the process in which a  $b\bar{b}$  pair is produced directly in the hard scattering. Although, due to interference, these three processes are mixed together, they can be cleanly separated in the soft/collinear limit. All current fixed-order programs with a massive  $b$  implement only FCR at NLO [79, 80], while GSP and FEX are only LO processes. This results in  $K$ -factors (NLO/LO) and renormalisation and factorisation scale dependence that are far larger than is expected from NLO calculations, as can be seen in

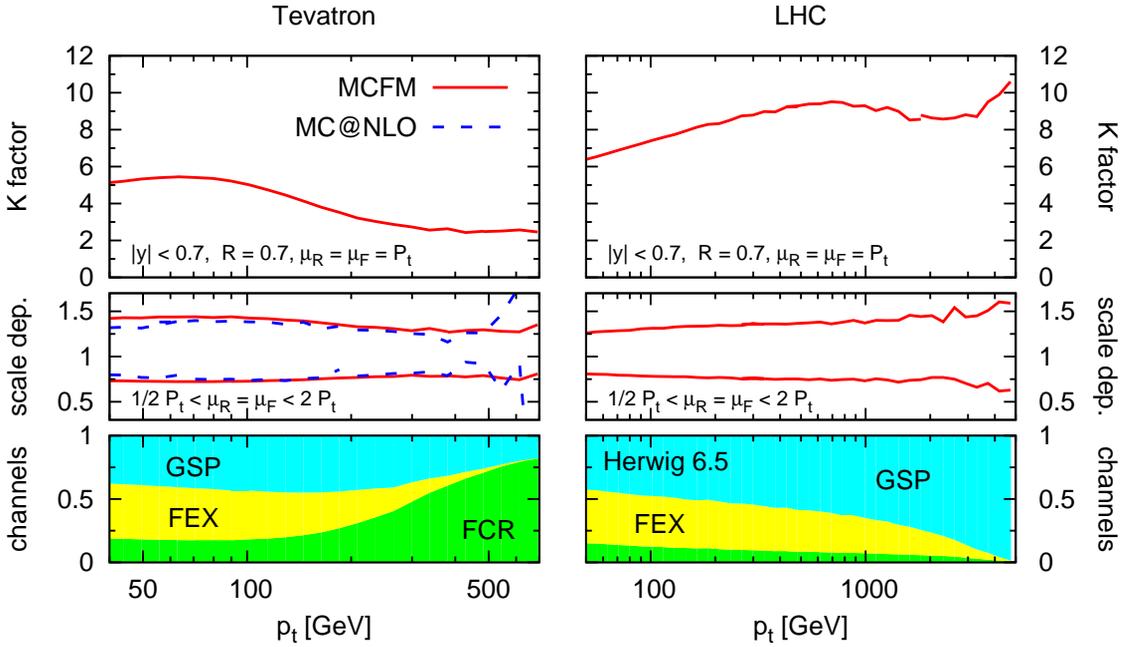


Fig. 8: Top:  $K$ -factors for inclusive  $b$ -jet spectrum as computed with MCFM, clustering particles into jets using the  $k_t$  jet-algorithm with  $R=0.7$ , and selecting jets in a central rapidity region ( $|y| < 0.7$ ). Middle: scale dependence obtained by simultaneously varying the renormalisation and factorisation scales by a factor two around  $p_t$ , the transverse momentum of the hardest jet in the event. For the Tevatron the scale uncertainty is computed also with MC@NLO. Bottom: breakdown of the HERWIG inclusive  $b$ -jet spectrum into the three major hard underlying channels contributions (for simplicity the small  $b\bar{b} \rightarrow b\bar{b}$  contribution is not shown).

fig. 8.<sup>13</sup> It is particularly instructive also to have a look at the bottom plots in the figure, which show the relevance of the various production channels as obtained from HERWIG [82]. Notice in particular how at the LHC GSP is the dominant process at any value of  $p_t$ . This is due to the fact that in  $pp$  collisions the process  $q\bar{q} \rightarrow b\bar{b}$ , the one responsible for FCR, is small also at high  $p_t$  due to the smallness of the antiquark distribution in the proton.

This situation can be significantly improved by exploiting an IRC safe definition of jet-flavour, such as the one outlined in the previous section. To overcome the experimental difficulty of discriminating  $b$  from  $\bar{b}$ , one can define a  $b$ -jet as a jet containing an odd number of  $b$ -hadrons without any risk for the IRC safety of the jet flavour [83]. In this case, the GSP contribution to  $b$ -jet production disappears immediately, because all jets with two  $b$ 's will be classified as gluon jets, and therefore will not contribute at all to  $b$ -jet cross sections. FEX contributions give rise to jets with a single  $b$ , so they cannot be eliminated by a jet-flavour algorithm. However, the FEX collinear logarithms are precisely those resummed in the  $b$  parton density, one of the ingredients of any PT calculation with massless  $b$ 's. Therefore one can compare experimental data for  $b$ -jet  $p_t$ -spectra obtained with the IRC definition of sec. 4.2 with PT predictions with

<sup>13</sup>Note that the addition of a parton shower as done in MC@NLO [81] does not solve the problem. This is because the underlying hard configurations remain the same as NLO, and have therefore the same collinear singularities.

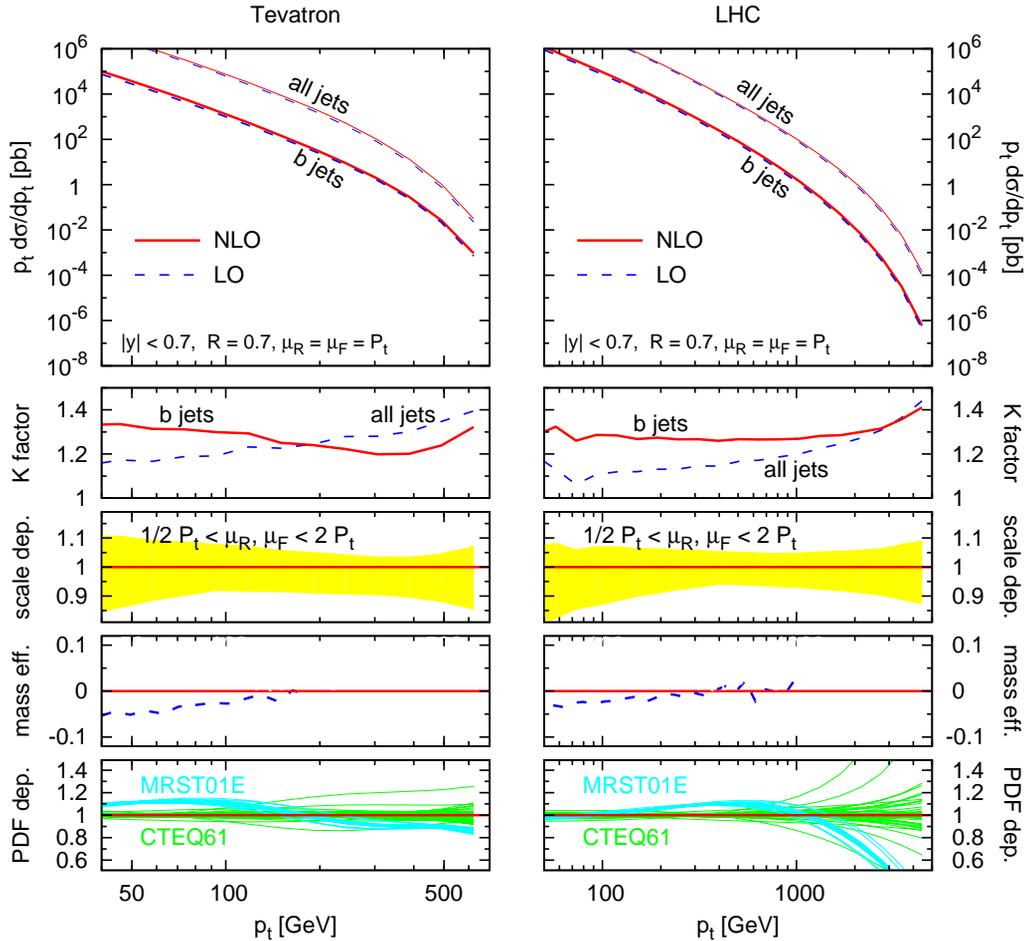


Fig. 9: The  $p_t$  spectrum for  $b$ -jets at the Tevatron (left) and at the LHC (right) obtained with NLO program NLOJET++. Below one can see, in order,  $K$ -factors for  $b$ -jets and all-flavoured inclusive jets, scale uncertainties obtained by varying *independently* renormalisation and factorisation scales by a factor of two, mass effects and PDF uncertainties.

massless  $b$ 's, which are available at NLO accuracy [84, 85]. Since all collinear singularities have been either eliminated or resummed, the difference between the massless and massive calculation should only involve powers of  $m_b^2/p_t^2$  (potentially enhanced by logarithms). The resulting NLO  $p_t$  spectra at the Tevatron and at the LHC are shown in fig. 9. There one can see that now the  $K$ -factors for  $b$ -jets are comparable to those for unflavoured jets, and moderate, indicating that the PT expansion is under control. Furthermore, scale uncertainties are at most 10%, and adding PDF uncertainties the overall theoretical error does not exceed 20%, except at very high  $p_t$  values at the LHC, where PDF's are less constrained. Note that mass effects are less than 5%, therefore not contributing significantly to the total uncertainty.

A technical difficulty to perform such a calculation is that no NLO program contains information on the flavour of produced partons. One is then forced to extract this information from

one's favourite NLO code (in our case NLOJET++ [85]). This procedure, although not straightforward, is nevertheless far easier than writing and testing a new code from scratch. Due to the relevance that jet-flavour algorithms can have for precision calculations we strongly encourage the authors of NLO codes to provide flavour information by default.

We remark that very similar results are obtained for charmed jet spectra. An interesting issue there is that predictions are very sensitive to possible intrinsic charm components of the proton [86], so that these observables can be exploited to set constraints on such intrinsic components.

A last remark concerns the feasibility of the experimental measurement of heavy flavour jets defined with our flavour algorithm. For a successful comparison between theory and experiment it is crucial to identify cases in which both heavy-flavoured particles are in the same jet, so as to label this jet as a gluon jet and eliminate the contribution of these configurations from the heavy-quark jet cross sections. Experimental techniques for double  $b$ -tagging in the same jet already exist [87] and steady progress is to be expected in the near future [88–90]. However one has always a limited efficiency for single  $b$  tagging, and even more for double  $b$ -tagging in the same jet. On the other hand preliminary studies indicate that one does not necessarily need high efficiencies, but what is more crucial is that one dominates the error on those efficiencies [83]. We look forward to further investigation in this direction.

**Acknowledgements.** This work has been done in collaboration with Gavin Salam.

## 5 Towards NNLO predictions for top quark production

*Author: M. Czakon*

Although discovered quite some time ago, the top quark has not been studied sufficiently to not deserve a special place in the LHC physics programme. This contribution to the workshop proceedings addresses part of the latter related to the top quark pair production cross section. While ideas of applications seem to have cristalized, there has also been progress in the evaluation of the next-to-next-to-leading order corrections. Here, I give some details of the methods.

The top quark has enjoyed a sustained attention for more than a decade since its discovery. Only this year, several theoretical studies have been published on its properties in view of the LHC. The interested reader is directed to [91]. A quantity of particular importance is the total production cross section. Without entering into a detailed discussion it is sufficient to say that one may expect a precision of measurement at the level of about 5% after a few years of LHC running, a number which on the one hand constitutes a challenge to the theory, and on the other opens the door for a few applications, of which only two will be mentioned here.

The first of the applications is indirect mass determination. Clearly, the total cross section is a decreasing function of the mass due mostly to the phase space dependence on the final states. A convenient representation of the connection between the error on the top quark mass,  $m_t$ , and the error on the total cross section,  $\sigma_{t\bar{t}}$ , is given by

$$\frac{\Delta\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}} \approx 5 \frac{\Delta m_t}{m_t}, \quad (24)$$

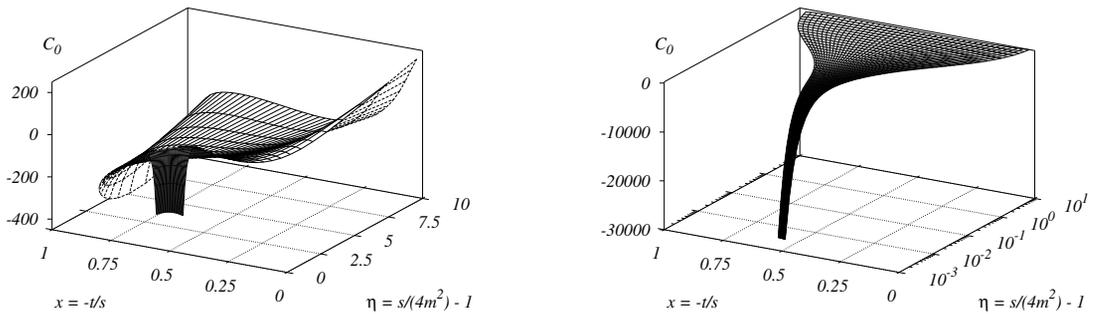


Fig. 10: Finite parts of the bosonic contributions to the two-loop amplitude in quark annihilation (most subleading color coefficient).

which is valid in a broad range around the current top quark mass. Clearly, this formula points at the possibility of determining  $m_t$  with an accuracy at the one percent level, as long as the theory prediction of  $\sigma_{t\bar{t}}$  is not a limiting factor. This is competitive with the best results from the Tevatron, but less precise than the ambitious goals of the LHC. The question remains, therefore, of the relevance of this method. A look at the way  $m_t$  is measured at present, and the variations of the central value implied, should convince a skeptic that it is important to have an independent measurement, which is far less sensitive on the kinematic reconstruction of hadronic final states.

The second application is gluon luminosity determination, to a large extent synonymous of the gluon PDF determination. While the standard luminosity monitor process for hadron colliders, Drell-Yan gauge boson production is sensitive mostly to quark PDFs, many of the non-standard processes and also the Higgs production process are induced by gluon fusion. A recent study by CTEQ [92], has shown that one can exploit the strong correlation between those cross sections and the top quark pair production cross section to reduce the errors. A prerequisite for success is a precision of 5% on both the theory and the experimental side.

In view of the above, a precise theory prediction for  $\sigma_{t\bar{t}}$  would be more than welcome. As far as fixed order perturbation theory is concerned, the result of [93] shows an error, judged by scale dependence, in excess of 10%. Since there is a substantial enhancement of the production rate due to soft gluon emission, one might expect that the knowledge of higher order corrections in the threshold regime would reduce the final uncertainty. This is indeed the case, as shown in various studies, of which the most recent are [94–96]. In the end, it is possible to obtain a prediction with a conservative error estimate slightly below 10%. While this number is not quite satisfactory, there is a second drawback to the approach based on threshold resummations. Namely, it does not fit a Monte-Carlo generator. With the high statistics of the LHC, MC programs are indispensable. All in all, it seems that having a fixed order result with next-to-next-leading accuracy would be a perfect solution. This statement is only strengthened by the fact, that the error from scale dependence induced would then amount to only 3% [94].

An NNLO prediction for a production process at the LHC needs four ingredients: 1) the two-loop virtual corrections, 2) the one-loop squared corrections, 3) the one-loop corrections with

an additional parton radiation, 4) the tree-level corrections with two additional partons radiated. Within the last one or two years, the first three points have been completed to a large extent for the case of  $\sigma_{t\bar{t}}$  [97–102]. Clearly, point 4) is trivial as long as all the partons are distinguishable. Performing the phase space integration over the unresolved configuration in 3) and 4) is the main remaining challenge. We are not going to discuss this issue, as it is not yet solved, but rather give a few details of the solution to point 1), which is an achievement in itself.

The main problem in the determination of the two-loop virtual corrections is the integration over the virtual momenta. The method adopted in [99] is based on a numerical solution of a system of differential equations [103]. It is suitable for problems with a relatively low number of scales and relies on the fact that Feynman integrals are smooth functions when evaluated above all thresholds as is here the case. The boundaries required are obtained from a series expansion solution to the differential equations around the high energy limit of the integrals derived in [97, 98]. While the integration of the system of equations is not fast enough to fit into a Monte-Carlo program, the presence of only two kinematic variables allows to use interpolation on a grid of precalculated values. The result for the most complicated color coefficient (most subleading term) in quark annihilation is shown in Fig. 10. The appropriate color decomposition is

$$\mathcal{A}^{(0,2)} = 2\text{Re} \langle \mathcal{M}^{(0)} | \mathcal{M}^{(2)} \rangle = 2(N^2 - 1) \quad (25)$$

$$\times \left( N^2 A + B + \frac{1}{N^2} C + N n_l D_l + N n_h D_h + \frac{n_l}{N} E_l + \frac{n_h}{N} E_h + n_l^2 F_l + n_l n_h F_{lh} + n_h^2 F_h \right).$$

The result for the gluon fusion channel is underway. While there are no new complications in the method itself, the number of integrals which need boundaries and have not been determined previously is about three times larger.

## 6 2- and 3-loop heavy flavor contributions to $F_2(x, Q^2)$ , $F_L(x, Q^2)$ and $g_{1,2}(x, Q^2)$

*Authors: I. Bierenbaum, J. Blümlein and S. Klein*

### 6.1 Introduction

In the case of single photon exchange, the deep-inelastic double differential scattering cross-section can be expressed in terms of the unpolarized structure functions  $F_2(x, Q^2)$  and  $F_L(x, Q^2)$ , and the polarized structure functions  $g_1(x, Q^2)$  and  $g_2(x, Q^2)$ . We are considering heavy flavor corrections to these functions. In the NLO approximation, the corrections were calculated semi-analytically in  $x$ -space for  $F_2(x, Q^2)$  and  $F_L(x, Q^2)$  in [104], with a fast implementation in Mellin  $N$ -space given in [105]. In the polarized case the NLO corrections are available only in the asymptotic case  $Q^2 \gg m^2$  [106, 107]. The  $c\bar{c}$ -contributions to these structure functions in the region of smaller values of Bjorken- $x$ , are of the order of 20-40 % and exhibit different scaling violations than the contributions due to massless partons, as shown in Figure 11. For the parameterization of the parton distribution functions we used [108]. Hence, a more precise determination of the parton distribution functions and the measurement of  $\Lambda_{\text{QCD}}$ , as reached in the non-singlet case [109], requires an extension of the heavy quark contributions to  $O(a_s^3)$ , as in the massless case, to perform the flavor-singlet analyzes consistently. This can be done by observing

that for  $Q^2 \gtrsim 10 m_c^2$ ,  $F_2^{c\bar{c}}(x, Q^2)$  is very well described by its asymptotic expression in the limit  $Q^2 \gg m^2$ , [110], where one can calculate the heavy flavor Wilson coefficients, the perturbative part of the structure functions, analytically. More precisely, the heavy flavor Wilson coefficients in the limit  $Q^2 \gg m^2$  are obtained as a convolution of the light-flavor Wilson coefficients with the corresponding massive operator matrix elements (OMEs) of flavor decomposed quarkonic and gluonic operators between massless parton states, which are obtained from the light-cone expansion. Here, we consider the level of twist-2 operators. The light Wilson coefficients are known up to three loops [111] and carry all the process dependence, whereas the OMEs, the objects to be calculated here, are universal and process-independent. Using this approximation, the heavy flavor Wilson coefficients are calculated for  $F_{2,L}^{c\bar{c}}(x, Q^2)$  to 2-loop order in [110, 112, 113] and for  $F_L^{c\bar{c}}(x, Q^2)$  to 3-loop order in [114]. First steps towards the asymptotic 3-loop corrections for  $F_2^{c\bar{c}}(x, Q^2)$  are made by the present authors by calculating the  $O(\varepsilon)$  terms of the 2-loop heavy operator matrix elements, [26, 115], contributing to the 3-loop heavy flavor Wilson coefficients via renormalization. The logarithmic contributions in  $(m^2/\mu^2)$  of the OMEs, as well as all pole terms in  $1/\varepsilon$ , are completely determined by renormalization, in this providing a check on the calculation, and containing in the single pole terms the respective contributions of the 3-loop anomalous dimensions. Furthermore, first steps towards a full 3-loop calculation of moments of the heavy flavor Wilson coefficients were undertaken. Here, the moments  $N = 2 \dots 12$  of the NNLO non-singlet (NS) and pure-singlet (PS) contributions of the OMEs were calculated. In addition, one obtains the corresponding contributions to the three-loop anomalous dimensions given in [116, 117], cf. also [118], which are confirmed in an independent calculation.

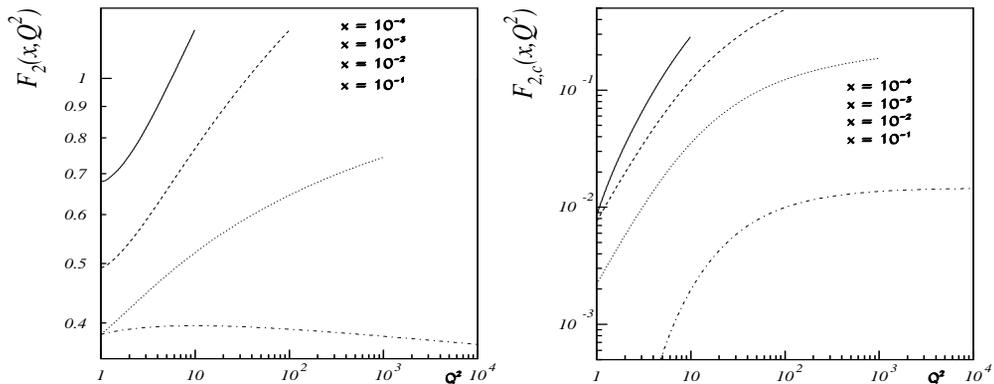


Fig. 11: The scaling violations of the light- and heavy-flavor contributions to the structure functions  $F_2^{\text{light}}$  and  $F_{2,c}$  at leading order.

## 6.2 Renormalization

Our calculation is done in Mellin space. The diagrams are of the self energy type with an additional operator insertion, which widely determines the dynamics and introduces the dependence on the Mellin variable  $N$ . The external particle is massless and on-shell. The scale is set by the mass of the heavy quark. After calculating the bare heavy flavor OMEs in  $D = 4 + \varepsilon$  dimensions and by using the Feynman-gauge, the renormalization is performed in four steps: We use

the on-shell scheme [119] for mass renormalization and the  $\overline{\text{MS}}$ -scheme for the charge renormalization.<sup>14</sup> The remaining two types of divergences, the UV and collinear singularities, are renormalized via the operator  $Z$ -factors and by mass factorization through the transition functions  $\Gamma$ . Denoting the completely unrenormalized OMEs by a double-hat,  $\hat{\hat{A}}$ , and those for which mass and coupling renormalization have already been performed by a single hat, the operator renormalization and mass factorization proceeds via

$$\mathbf{A} = \mathbf{Z}^{-1} \hat{\mathbf{A}} \mathbf{\Gamma}^{-1}, \quad (26)$$

which constitutes a matrix equation in the singlet case. This equation allows us to predict the pole-structure of the OMEs under consideration. The  $Z$ -factors read

$$\begin{aligned} Z_{ij}(N, a_s, \varepsilon) = & \delta_{i,j} + a_s S_\varepsilon \frac{\gamma_{ij,0}}{\varepsilon} + a_s^2 S_\varepsilon^2 \left\{ \frac{1}{\varepsilon^2} \left[ \frac{1}{2} \gamma_{im,0} \gamma_{mj,0} + \beta_0 \gamma_{ij,0} \right] + \frac{1}{2\varepsilon} \gamma_{ij,1} \right\} \\ & + a_s^3 S_\varepsilon^3 \left\{ \frac{1}{\varepsilon^3} \left[ \frac{1}{6} \gamma_{in,0} \gamma_{nm,0} \gamma_{mj,0} + \beta_0 \gamma_{im,0} \gamma_{mj,0} + \frac{4}{3} \beta_0^2 \gamma_{ij,0} \right] \right. \\ & \left. + \frac{1}{\varepsilon^2} \left[ \frac{1}{6} (\gamma_{im,1} \gamma_{mj,0} + 2\gamma_{im,0} \gamma_{mj,1}) + \frac{2}{3} (\beta_0 \gamma_{ij,1} + \beta_1 \gamma_{ij,0}) \right] + \frac{\gamma_{ij,2}}{3\varepsilon} \right\} \end{aligned} \quad (27)$$

They are related to the anomalous dimensions of the twist-2 operators via  $\gamma = \mu \partial \ln \mathbf{Z}(\mu) / \partial \mu$ , allowing to express them in terms of the anomalous dimensions up to an arbitrary order in the strong coupling constant  $a_s := \alpha_s / (4\pi)$  (cf. [115] up to  $O(a_s^3)$ ). Additionally, we would have  $\mathbf{\Gamma} = \mathbf{Z}^{-1}$ , if all quark lines were massless, which, however, has to be modified here since we always have at least one heavy quark line. From these equations, one can infer that for operator renormalization and mass factorization at  $O(a_s^3)$ , the anomalous dimensions up to NNLO, [116, 117], together with the 1-loop heavy OMEs up to  $O(\varepsilon^2)$  and the 2-loop heavy OMEs up to  $O(\varepsilon)$  are needed. The last two quantities enter since they multiply  $Z$ - and  $\Gamma$ -factors containing poles in  $\varepsilon$  (cf. [115]).

To see this in more detail, let us consider as an example the term  $A_{gq,Q}$ , which emerges for the first time at  $O(a_s^2)$ . By applying Eq. (26), one obtains at  $O(a_s^2)$  the renormalized OME

$$A_{gq,Q}^{(2)} = \hat{\hat{A}}_{gq}^{(2)} + Z_{gq}^{-1,(2)} + \left( Z_{gg}^{-1,(1)} + \hat{\hat{A}}_{gg,Q}^{(1)} \right) \mathbf{\Gamma}_{gq}^{-1,(1)}.$$

Here, the term  $\hat{\hat{A}}_{gq,Q}^{(1)}$ , cf. [120], enters through mixing. Note that since we consider only terms involving at least one heavy quark, we adopt the definition  $\hat{\gamma} := \gamma(n_f + 1) - \gamma(n_f)$  for the anomalous dimensions in order to obtain the correct color projection. Now we can predict the structure of the unrenormalized result to be

$$\hat{\hat{A}}_{gq,Q}^{(2)} = \left( \frac{m^2}{\mu^2} \right)^\varepsilon \left[ \frac{2\beta_{0,Q}}{\varepsilon^2} \gamma_{gq}^{(0)} + \frac{\hat{\gamma}_{gq}^{(1)}}{2\varepsilon} + a_{gq,Q}^{(2)} + \varepsilon \bar{a}_{gq,Q}^{(2)} \right], \quad (28)$$

where we see the LO and NLO anomalous dimensions and  $\beta_{0,Q} = -(4/3)T_F$  occurring in the pole terms. The terms which are in general not predictable are the constant and  $O(\varepsilon)$ -terms,

<sup>14</sup>For the latter we make the requirement that the heavy quark loop contributions to the gluon self-energy,  $\Pi(p^2, m^2)$ , are renormalized in such a way that  $\Pi(0, m^2) = 0$ , cf. [110, 112, 113, 115].

which, however, enter the pole and constant terms of a 3-loop OME, as mentioned above. In this particular case here, the calculation in Mellin-space in terms of Feynman-parameters is straightforward, cf. [112, 113], and a representation in Euler- $\Gamma$  functions can be obtained even to all orders in  $\varepsilon$ , where we reproduced the pole terms of Eq. (28), [121]. As a last remark, note that we consider charm quark contributions here, while for heavier quarks decoupling [122] has to be applied.

### 6.3 $O(\varepsilon)$ at 2-loops

The appearance of the constant and  $O(\varepsilon)$  terms in the renormalization process of the OMEs has been worked out in some detail in Ref. [115], [123], where we presented the  $O(\varepsilon)$  terms  $\bar{a}_{Qg}^{(2)}$ ,  $\bar{a}_{qq,Q}^{(2),NS}$  and  $\bar{a}_{Qq}^{(2),PS}$  in the unpolarized case. The term  $\bar{a}_{gg,Q}^{(2)}$  was given in [26]. The last missing 2-loop  $O(\varepsilon)$  term corresponds to the heavy OME  $A_{gg,Q}^{(2)}$ , [107, 121]. The corresponding constant contribution was calculated before in Ref. [120]. It contributes through operator mixing to the  $T_F^2$ -term of  $A_{Qq}^{(3),PS}$ , which we consider in this paper.

Since we perform our calculation in Mellin space, all results are given in terms of harmonic sums, [124, 125], the argument of which we have set equal to  $N$ . Thus, the results of the constant and  $O(\varepsilon)$ -terms of the above-mentioned  $A_{gg,Q}^{(2)}$ , for example, are given by:

$$a_{qq,Q}^{(2)} = T_F C_F \left\{ \frac{4}{3} \frac{N^2 + N + 2}{(N-1)N(N+1)} (S_2 + S_1^2 + 2\zeta_2) - \frac{8}{9} \frac{8N^3 + 13N^2 + 27N + 16}{(N-1)N(N+1)^2} S_1 + \frac{8}{27} \frac{P_1}{(N-1)N(N+1)^3} \right\}, \quad (29)$$

$$\bar{a}_{qq,Q}^{(2)} = T_F C_F \left\{ \frac{2}{9} \frac{N^2 + N + 2}{(N-1)N(N+1)} (-2S_3 - 3S_2 S_1 - S_1^3 + 4\zeta_3 - 6\zeta_2 S_1) + \frac{2}{9} \frac{8N^3 + 13N^2 + 27N + 16}{(N-1)N(N+1)^2} (2\zeta_2 + S_2 + S_1^2) - \frac{4P_1 S_1}{27(N-1)N(N+1)^3} + \frac{4P_2}{81(N-1)N(N+1)^4} \right\}, \quad (30)$$

$$P_1 = 43N^4 + 105N^3 + 224N^2 + 230N + 86.$$

$$P_2 = 248N^5 + 863N^4 + 1927N^3 + 2582N^2 + 1820N + 496.$$

The representation in Mellin-space allowed us to use various analytic and algebraic relations between harmonic sums, [126–128], to obtain a more compact result. Together with the result of Eq. (30), all 2-loop  $O(\varepsilon)$  terms of the heavy OMEs in the unpolarized case are known by now. A corresponding calculation has been performed for the polarized case up to  $O(\varepsilon)$  [107] extending the results of Ref. [106]. The contributions to the structure function  $g_2(x, Q^2)$  can be obtained using Wandzura-Wilczek relations, cf. [129, 130]. For the respective formulae we refer to the original paper.

### 6.4 Fixed moments at 3–loops

We start by calculating the diagrams for fixed even values of Mellin  $N$ . At this order, new operator vertices appear with three and four gluonic lines, for which the Feynman–rules had not yet been derived before. The necessary 3–loop diagrams are generated using QGRAF [131] and are genuinely given as tensor integrals due to the operators contracted with the light–cone vector  $\Delta$ ,  $\Delta^2 = 0$ . The calculation proceeds in the following steps: first, the contraction with the light–cone vector is made undone, which leaves tensor integrals for each diagram. For each value of Mellin  $N$  under consideration, one then constructs a projector, which, applied to the tensor integrals, projects onto the desired  $N$ . We consider  $N = 2, \dots, 12$ . The color factors of the diagrams are calculated using [132]. A generalization to higher moments is straightforward, however, the computing time increases rapidly. The diagrams are then translated into a form, which is suitable for the program MATAD [133], doing the expansion in  $\varepsilon$  for the corresponding massive three–loop tadpole–type diagrams. We have implemented all these steps into a FORM–program, cf. [134], and tested it against various two–loop results, including the result for  $\hat{A}_{gq,Q}^{(2)}$ , Eq. (28), and found agreement.

The first 3–loop objects we are investigating are the OMEs  $A_{qq,Q}^{\text{NS}}$ , cf. [121], and  $A_{Qq}^{\text{PS}}$ . All diagrams contain two inner quark loops, where the quark to which the operator insertion couples is heavy and the other one may be heavy or light. The latter two cases can be distinguished by a factor  $n_f$ , denoting the number of light flavors, in the result. From Eq. (26), we can obtain the pole structure of the the  $T_F^2$  terms of the completely unrenormalized PS OME:

$$\hat{A}_{Qq}^{(3),\text{PS}} \Big|_{T_F^2} = \left( \frac{m^2}{\mu^2} \right)^{3\varepsilon/2} \left\{ 2 \frac{n_f + 4}{3\varepsilon^3} \beta_{0,Q} \hat{\gamma}_{qg}^{(0)} \gamma_{gq}^{(0)} + \frac{1}{\varepsilon^2} \left( \frac{2 - n_f}{6} \hat{\gamma}_{qg}^{(0)} \hat{\gamma}_{gq}^{(1)} - (n_f + 1) \frac{4}{3} \beta_{0,Q} \hat{\gamma}_{\text{PS}}^{(1)} \right) + \frac{1}{\varepsilon} \left( \frac{n_f + 1}{3} \hat{\gamma}_{\text{PS}}^{(2)} - 4(n_f + 1) \beta_{0,Q} a_{Qq}^{(2),\text{PS}} - n_f \frac{\zeta_2 \beta_{0,Q}}{4} \hat{\gamma}_{qg}^{(0)} \gamma_{gq}^{(0)} + \hat{\gamma}_{qg}^{(0)} a_{gq,Q}^{(2)} \right) + a_{Qq}^{(3),\text{PS}} \Big|_{T_F^2} \right\} \quad (31)$$

The  $n_f$  dependence is written explicitly and  $\hat{\gamma}_{\text{PS}}^{(2)}$  is the term  $\propto n_f^2$  of the NNLO anomalous dimension  $\gamma_{\text{PS}}^{(2)}$ . It is not possible to factor out  $(n_f + 1)$ , not even in the triple pole term. This is due to the interplay of the prescription for coupling constant renormalization we have adopted and the fact that the transition functions  $\Gamma$  apply to sub–graphs containing massless lines only. We have calculated the above term using MATAD for  $N = 2, \dots, 12$  and all pole terms agree with Eq. (31). Detailed Tables of these results can be found in [121] and a further upcoming paper. Using Eqs. (29,31), one can obtain moments for the 3–loop anomalous dimension  $\gamma_{\text{PS}}^{(2)}|_{T_F^2}$ , see also [121] and a corresponding paper in preparation. These latter results agree with the results from [117]. Here one has to make the replacement  $n_f \rightarrow n_f(2T_F)$ , with  $T_F = 1/2$ , and to multiply by 2, to account for the different convention for the  $Z$ –factors we adopted. As an example consider the renormalized result for the second moment. Applying Eq. (26), we obtain

$$A_{Qq}^{(3),\text{PS}} \Big|_{N=2,T_F^2} = C_F T_F^2 \left\{ -\frac{128}{81} \ln^3 \left( \frac{m^2}{\mu^2} \right) - \frac{32}{27} \ln^2 \left( \frac{m^2}{\mu^2} \right) - \frac{5344}{243} \ln \left( \frac{m^2}{\mu^2} \right) + \frac{53144}{2187} - \frac{3584}{81} \zeta_3 + n_f \left( -\frac{128}{81} \ln^3 \left( \frac{m^2}{\mu^2} \right) + \frac{32}{27} \ln^2 \left( \frac{m^2}{\mu^2} \right) - \frac{5104}{243} \ln \left( \frac{m^2}{\mu^2} \right) - \frac{34312}{2187} + \frac{1024}{81} \zeta_3 \right) \right\} \quad (32)$$

As in Eq. (32), we observe for all moments in the NS and PS case that the terms  $\propto \zeta_2$  disappear after renormalization, since the corresponding terms in the light flavor Wilson coefficients do not contain even  $\zeta$ -values. This provides us with a further check on our calculation, since it is a general observation made in many  $D = 4$  calculations.

For the  $T_F^2$ -terms of the heavy OME  $A_{qq,Q}^{(3),NS}$ , a formula similar to Eq. (31) can be derived, cf. [121]. Using again MATAD, we have calculated the first 6 non-vanishing moments of the completely unrenormalized expression. The pole terms we obtain agree with what one expects from Eq. (26) and after renormalization, we again observe that there are no  $\zeta_2$ 's left anymore. Additionally, the values for the moments of the terms  $\propto T_F$  in  $\gamma_{NS}^{(2)}$  agree with those in Refs. [116–118].

## 6.5 Conclusions and outlook

All  $O(\varepsilon)$  contributions to the unpolarized and most of the polarized heavy quark OMEs for general Mellin variable  $N$  at  $O(a_s^2)$  were calculated which are needed for the renormalization at  $O(a_s^3)$ . This part of the calculation makes significant use of the representation of Feynman-integrals in terms of generalized hypergeometric and related functions, omitting the integration-by-parts method. The solution of the sums beyond those which could be performed by `summer` [125], required new techniques and were solved using `SIGMA` [135]. Concerning the structure of the result, we find the *universal* pattern as observed in case of the massless 2-loop Wilson coefficients and related quantities in terms of harmonic sums [126, 127, 136–138]. Furthermore, we installed a program chain to calculate the corresponding 3-loop diagrams to  $O(a_s^3)$  using MATAD. As a first step, we obtained the moments of the heavy OMEs  $\hat{A}_{qq,Q}^{(3),NS}$  and  $\hat{A}_{Qq}^{(3),PS}$ , for which we found agreement with the general pole structure expected from renormalization. This provides us with a good check on the method we apply for our calculation. For the calculation of high moments we will apply `TFORM`, [139], in the future. In the same way all other contributions to the heavy quark OMEs will be calculated.

**Acknowledgments.** We would like to thank M. Steinhauser and J. Vermaseren for useful discussions and M. Steinhauser for a `FORM 3.0` compatible form of the code MATAD.

## 7 Heavy quark and quarkonium production in the Regge limit of QCD

*Author: V. Saleev*

We study production of hadrons containing charm and beauty quarks at HERA and Tevatron Colliders in the framework of the quasi-multi-Regge-kinematics approach at leading order in the strong-coupling constant  $\alpha_s$ . To describe heavy quark hadronization we use the fragmentation approach in case of  $D$ - and  $B$ -meson production, or the factorization formalism of nonrelativistic QCD at leading order in the relative velocity  $v$  of heavy quarks in quarkonia in case of heavy quarkonium production.

## 7.1 Theoretical basis

Heavy quark and quarkonium production at high energies has provided a useful laboratory for testing the perturbative quantum chromodynamics (pQCD) as well as the interplay of perturbative and nonperturbative phenomena in QCD. Also these studies are our potential for the observation of a new dynamical regime, namely the high-energy Regge limit, which is characterized by the following condition  $\sqrt{S} \gg \mu \gg \Lambda_{QCD}$ , where  $\sqrt{S}$  is the total collision energy in the center of mass reference frame,  $\Lambda_{QCD}$  is the asymptotic scale parameter of QCD,  $\mu$  is the typical energy scale of a hard interaction.

The phenomenology of strong interactions at high energies exhibits a dominant role of gluon interactions in heavy quark and quarkonium production. In the conventional parton model [140], the initial-state gluon dynamics is controlled by the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution equations [141], in which it is assumed that  $S > \mu^2 \gg \Lambda_{QCD}^2$ . Thus, the DGLAP evolution equation takes into account only one large logarithm, namely  $\ln(\mu/\Lambda_{QCD})$  and the collinear approximation is used, in which the transverse momenta of the initial gluons ( $k_T$ ) are neglected.

In the Regge limit the summation of large logarithms  $\ln(\sqrt{S}/\mu)$  in the evolution equation can then be more important than the one of the  $\ln(\mu/\Lambda_{QCD})$  terms. In this case, the non-collinear gluon dynamics is described by the Balitsky-Fadin-Kuraev-Lipatov (BFKL) evolution equation [142]. In the region under consideration, the transverse momenta of the incoming gluons and their off-shell properties can no longer be neglected, and we deal with *Reggeized* gluons. As the theoretical framework for this kind of high-energy phenomenology, the quasi-multi-Regge-kinematics (QMRK) approach [143], which is based on the effective quantum field theory implemented with the non-abelian gauge-invariant action [144], can be used. The *Reggeization* of particles or amplitudes is the well-known effect for electrons in high-energy quantum electrodynamics (QED) [145] and for gluons and quarks in QCD [142, 146]. Roughly speaking, the *Reggeization* is a trick, which gives an opportunity to take into account efficiently large radiative corrections to the processes under Regge limit condition beyond the collinear approximation. The main ingredients of the QMRK approach are the effective vertices of Reggeon-Reggeon-Particle (RRP) or Reggeon-Particle-Particle (RPP) interactions, which can be obtained from the effective action [144].

The factorization formalism of nonrelativistic QCD (NRQCD) [147] is a theoretical framework for the description of heavy-quarkonium production and decay. The factorization hypothesis of NRQCD assumes the separation of the effects of long and short distances in heavy-quarkonium production. NRQCD is organized as a perturbative expansion in two small parameters, the strong-coupling constant  $\alpha_s$  and the relative velocity  $v$  of heavy quarks in quarkonium.

The studies of the open heavy-flavour production at high energies show that in calculations the precise implementation of the effect of heavy quark fragmentation is needed to describe data [45–47, 148]. The approach used here applies the universal fragmentation functions (FFs) [45–47], which satisfy DGLAP evolution equations and are fitted to  $e^+e^-$  annihilation data for the open heavy-flavour production from CERN LEP1.

Both models, the NRQCD and the fragmentation approach, don't depend on the choice of high-energy factorization scheme and they can be used in calculations both in the conventional

collinear parton model and in the QMRK approach.

## 7.2 Charmonium production at Tevatron and HERA

During the last decade, the CDF Collaboration at the Tevatron [149, 150] collected data on charmonium production at the energies  $\sqrt{S} = 1.8$  TeV (run I) and  $\sqrt{S} = 1.96$  TeV (run II) in the central region of pseudorapidity  $|\eta| < 0.6$ . In contrast to previous analysis in the collinear parton model [151] or the  $k_T$ -factorization approach [152–154], we perform a joint fit to the run-I and run-II CDF data [149, 150] to obtain the color-octet nonperturbative matrix elements (NMEs) for  $J/\psi$ ,  $\chi_{cJ}$ , and  $\psi'$  mesons. The run-II data include region of small  $J/\psi$  transverse momentum, which can't be described principally in the collinear parton model, but this region is important for fit procedure. Our calculations [155, 156] are based on exact analytical expressions for the relevant *Reggeized* amplitudes, which were previously unknown in the literature ( $R + R \rightarrow H$ ,  $R + R \rightarrow H + g$ , and  $R + P \rightarrow H$ , where  $H$  is  $q\bar{q}$ -pair in the fixed quantum state,  $R$  is the *Reggeized* gluon). Our fits include five experimental data sets, which come as  $p_T$  distributions of  $J/\psi$  mesons from direct production, prompt production,  $\chi_{cJ}$  decays, and  $\psi'$  decays in run I, and from prompt production in run II. In the Table I of Ref. [155, 156], we present our fit results for the relevant color-octet NMEs for three different choices of unintegrated gluon distribution function, namely JB [157], JS [158], and KMR [159]. Our fits to the Tevatron data turned out to be satisfactory, except for the one to the  $\chi_{cJ}$  sample based on the JB gluon density in the proton, where the fit result significantly exceeded the measured cross section in the small- $p_T$  region, as it is shown in Figs. 4-5 of Ref. [155, 156]. We see also that color-octet contribution in case of  $\chi_{cJ}$  production is being quite unimportant. Considering the color-octet NMEs relevant for the  $J/\psi$ ,  $\psi'$  and  $\chi_{cJ}$  production mechanisms, we can formulate the following heuristic rule for favored transitions from color-octet to color-singlet states:  $\Delta L \simeq 0$  and  $\Delta S \simeq 0$ ; *i.e.*, these transitions are doubly chromoelectric and preserve the orbital angular momentum and the spin of the heavy-quark bound state.

At HERA, the cross section of prompt  $J/\psi$  production was measured in a wide range of the kinematic variables both in photoproduction [160], at small values of photon virtuality  $Q^2$ , and deep-inelastic scattering (DIS) [161], at large values of  $Q^2$ . In the Figs. 6-9 of Ref. [155, 156], our NRQCD predictions in the high-energy factorization approach, evaluated with the NMEs from Table 1 of Ref. [155, 156], are compared with the HERA data [160, 161]. In this regime, where the contribution of  $2 \rightarrow 1$  subprocesses is suppressed, the LO NRQCD predictions in the QMRK approach are mainly due to the color-singlet channels and are therefore fairly independent of the color-octet NMEs. Thus, our results agree well with the data and with the previous calculations in the color singlet model (CSM) [162], up to minor differences in the choice of the color-singlet NMEs and the  $c$ -quark mass. Let us note that first theoretical prediction for  $J/\psi$  photoproduction in the CSM and the  $k_T$ -factorization scheme has been done 15 years ago in Ref. [163].

## 7.3 Bottomonium production at the Tevatron

The CDF Collaboration measured the  $p_T$  distributions of  $\Upsilon(1S)$ ,  $\Upsilon(2S)$ , and  $\Upsilon(3S)$  mesons in the central region of rapidity ( $y$ ),  $|y| < 0.4$ , at  $\sqrt{S} = 1.8$  TeV (run I) [164] and that of the  $\Upsilon(1S)$  meson in the rapidity regions  $|y| < 0.6$ ,  $0.6 < |y| < 1.2$ , and  $1.2 < |y| < 1.8$  at  $\sqrt{S} = 1.96$  TeV (run II) [165]. In both cases, the  $S$ -wave bottomonia were produced promptly, *i.e.*, directly or

via non-forbidden decays of higher-lying  $S$ - and  $P$ -wave bottomonium states, including cascade transitions such as  $\Upsilon(3S) \rightarrow \chi_{b1}(2P) \rightarrow \Upsilon(1S)$ .

In contrast to previous analysis in the collinear parton model [166], we perform a joint fit to the CDF data from run I [164] and run II [165] for all  $p_T$  values, including the small- $p_T$  region. Comparing the color-singlet and color-octet contributions, we observe that the latter is dominant in the  $\Upsilon(3S)$  case and in the  $\Upsilon(2S)$  case for  $p_T \geq 13$  GeV, while it is of minor importance in the  $\Upsilon(1S)$  case in the whole  $p_T$  range considered. The fits based on the KMR, JB, and JS gluons turned out to be excellent, fair, and poor, respectively. They yielded small to vanishing values for the color-octet NMEs, see Table II of Ref. [167], especially when the estimated feed-down contributions from the as-yet unobserved  $\chi_{bJ}(3P)$  states were included. The presented analysis in Ref. [167], together with the investigation of charmonium production [155, 156], suggest that the color-octet NMEs of bottomonium are more strongly suppressed than those of charmonium as expected from the velocity scaling rules of NRQCD.

Using obtained NMEs for bottomonium and charmonium states we have done predictions for the LHC Collider at the energy  $\sqrt{S} = 14$  TeV, which are presented in Figs. 14-17 of Ref. [168].

#### 7.4 Open heavy-flavour production at HERA and Tevatron

At HERA  $D$ -meson production has been studied both in the photo-production processes and in the deep inelastic scattering (DIS) processes. The data are presented by H1 and ZEUS Collaborations for different spectra, see Refs. [169, 170]. The lowest order in  $\alpha_s$  processes of heavy quark photoproduction or electroproduction in the QMRK approach in the massive  $c$ -quark scheme are the following:  $\gamma(\gamma^*) + R \rightarrow c + \bar{c}$  - direct production and  $R_\gamma + R \rightarrow c + \bar{c}$  - resolved production, where  $R$  is the *Reggeized* gluon from a proton or  $R_\gamma$  is the one from a photon.

We find approximate agreement of our results with data from HERA for  $p_T$  spectra of  $D^*$ -meson production, the pseudo-rapidity spectra are described well only at the large  $p_T \geq 6$  GeV, see Figs. 3-6 in Ref. [171]. These conclusions are true both for photoproduction and for  $D^*$  production in DIS.

Recently the CDF Collaboration measured the differential cross sections  $d\sigma/dp_T$  for the inclusive production of  $D^0$ ,  $D^+$ ,  $D^{*+}$ , and  $D_s^+$  mesons [172] in  $p\bar{p}$  collisions at the Fermilab Tevatron (run I and run II) as functions of transverse momentum ( $p_T$ ) in the central rapidity ( $y$ ) region. At the LO QMRK approach the parton subprocesses for heavy quark production in hadron collisions are:  $R + R \rightarrow c + \bar{c}$  and  $Q + \bar{Q} \rightarrow c + \bar{c}$ , where  $Q$  is the *Reggeized* quark in a proton. The squared matrix elements of all above mentioned processes, excluding last one with *Reggeized* initial quarks, are known in the literature [143, 173, 174]. The contribution of the subprocess  $Q + \bar{Q} \rightarrow c + \bar{c}$  is studied for the first time [175].

In the paper [176], we explored the usefulness of the quark-Reggeization hypothesis in the framework of the QMRK approach by studying several observables of inclusive charm production at LO, namely the charm structure function  $F_{2,c}$  of the proton measured at HERA as well as the one-particle-inclusive cross sections of  $D^{*\pm}$  and  $D_s^\pm$  photoproduction in  $ep$  collisions at HERA and of  $D^0$ ,  $D^\pm$ ,  $D^{*\pm}$ , and  $D_s^\pm$  hadroproduction in  $p\bar{p}$  collisions at the Tevatron Collider. In all three cases, we found satisfactory agreement between our default predictions and the ex-

perimental data, which is quite encouraging in view of the simplicity of our LO expressions for the partonic cross sections. By contrast, in the collinear parton model of QCD, the inclusion of NLO corrections is necessary to achieve such a degree of agreement. We thus recover the notion that the QMRK approach is a powerful tool for the theoretical description of QCD processes in the high-energy limit and automatically accommodates an important class of corrections that lie beyond the reach of the collinear parton model at LO.

The first theoretical prediction for the beauty production at Tevatron [177] based on high-energy factorization scheme and Reggeon-Reggeon effective vertices [143] for the process  $R + R \rightarrow b + \bar{b}$  has been done in Ref. [178]. It was shown that both  $p_T$ -spectra and total cross section of  $B$ -mesons can be described well with KMS unintegrated gluon distribution function [179]. We performed these calculations with KMR [159] unintegrated distribution functions and Peterson  $b$ -quark fragmentation function [180], and have found good agreement with data too. Thus, in case of  $b$ -quark production, contrary to  $c$ -quark production, theoretical description of data both for  $B$ -mesons and for bottomonia looks well grounded and more simple. The  $c$ -quark mass is not large enough and nonperturbative effects in the hadronization of  $c$ -quarks need more careful description.

## 7.5 Conclusions

Our results show that the QMRK approach is a very powerful tool in the high-energy phenomenology of heavy quark and quarkonium production. Of course, there is a number of non-solved problems yet, such as the correct description of  $J/\psi$  polarization [181] and an estimation of NLO corrections for relevant processes. At the LHC Collider the conditions of application of the QMRK approach for heavy quark production will be satisfied with higher accuracy, therefore we see many future applications of this approach in a new kinematic regime.

The author thanks B. Kniehl, D. Vasin and A. Shipilova for cooperation in study of presented results. We thank also L. Lipatov, M. Ryskin, G. Kramer, H. Spiesberger and O. Teryaev for useful discussions.

## 8 Upsilononium polarization as a touchstone in understanding the parton dynamics in QCD

*Authors: S. Baranov and N. Zotov*

Nowadays, the production of heavy quarkonium states at high energies is under intense theoretical and experimental study [182, 183]. The production mechanism involves the physics of both short and long distances, and so, appeals to both perturbative and nonperturbative methods of QCD. This feature gives rise to two competing theoretical approaches known in the literature as the color-singlet and color-octet models. According to the color-singlet approach, the formation of a colorless final state takes place already at the level of the hard partonic subprocess (which includes the emission of hard gluons when necessary). In the color-octet model, also known as nonrelativistic QCD (NRQCD), the formation of a meson starts from a color-octet  $Q\bar{Q}$  pair and proceeds via the emission of soft nonperturbative gluons.

Originally, the color-octet model was introduced to overcome the discrepancy between the large  $J/\psi$  production cross section measured in  $pp$  interactions at the Tevatron and the results of theoretical calculations based on the standard perturbative QCD. The problem was apparently

solved by attributing the discrepancy to the hypothetical contributions from the intermediate color-octet states, which must obey certain hierarchy in powers of the relative velocity of the quarks in a bound system. However, the numerical estimates of these contributions extracted from the analysis of Tevatron data are at odds with the HERA data, especially as far as the inelasticity parameter  $z = E_\psi/E_\gamma$  is concerned [184]. In the  $k_t$ -factorization approach, the values of the color-octet contributions obtained as fits of the Tevatron data appear to be substantially smaller than the ones in the collinear scheme, or even can be neglected at all [153, 155, 185, 186].

The first attempts to solve the quarkonium polarization problem within the  $k_t$ -factorization approach were made in the pioneering work [187] (see also [188]) for  $ep$  collisions and in Refs. [154, 185] for  $pp$  collisions. It was emphasised that the off-shellness of the initial gluons, the intrinsic feature of the  $k_t$ -factorization approach, has an immediate consequence in the longitudinal polarization of the final state  $J/\psi$  mesons.

The goal of this paper is to derive theoretical predictions on the polarization of  $\Upsilon$  mesons produced at the Fermilab Tevatron and CERN LHC. In the  $k_t$ -factorization approach, the cross section of a physical process is calculated as a convolution of the partonic cross section  $\hat{\sigma}$  and the unintegrated parton distribution  $\mathcal{F}_g(x, k_T^2, \mu^2)$ , which depend on both the longitudinal momentum fraction  $x$  and transverse momentum  $k_T$ :

$$\sigma_{pp} = \int \mathcal{F}_g(x_1, k_{1T}^2, \mu^2) \mathcal{F}_g(x_2, k_{2T}^2, \mu^2) \hat{\sigma}_{gg}(x_1, x_2, k_{1T}^2, k_{2T}^2, \dots) dx_1 dx_2 dk_{1T}^2 dk_{2T}^2. \quad (33)$$

In accordance with [173, 189–191], the off-shell gluon spin density matrix is taken in the form

$$\overline{\epsilon_g^\mu \epsilon_g^{*\nu}} = p_p^\mu p_p^\nu x_g^2 / |k_T|^2 = k_T^\mu k_T^\nu / |k_T|^2. \quad (34)$$

In all other respects, our calculations follow the standard Feynman rules.

In order to estimate the degree of theoretical uncertainty connected with the choice of unintegrated gluon density, we use two different parametrizations, which are known to show the largest difference with each other, namely, the ones proposed in Refs. [189, 191] and [192]. In the first case [189], the unintegrated gluon density is derived from the ordinary (collinear) density  $G(x, \mu^2)$  by differentiating it with respect to  $\mu^2$  and setting  $\mu^2 = k_T^2$ . Here we use the LO GRV set [193] as the input collinear density. In the following, this will be referred to as dGRV parametrisation. The other unintegrated gluon density [192] is obtained as a solution of leading order BFKL equation [191] in the double-logarithm approximation. Technically, it is calculated as the convolution of the ordinary gluon density with some universal weight factor. This will be referred to as JB parametrisation.

The production of  $\Upsilon$  mesons in  $pp$  collisions can proceed via either direct gluon-gluon fusion or the production of  $P$ -wave states  $\chi_b$  followed by their radiative decays  $\chi_b \rightarrow \Upsilon + \gamma$ . The direct mechanism corresponds to the partonic subprocess  $g + g \rightarrow \Upsilon + g$  which includes the emission of an additional hard gluon in the final state. The production of  $P$ -wave mesons is given by  $g + g \rightarrow \chi_b$ , and there is no emission of any additional gluons. All the other parameters are the same as in our previous paper [194].

The polarization state of a vector meson is characterized by the spin alignment parameter  $\alpha$  which is defined as a function of any kinematic variable as  $\alpha(\mathcal{P}) = (d\sigma/d\mathcal{P} - 3d\sigma_L/d\mathcal{P}) / (d\sigma/d\mathcal{P} +$

$d\sigma_L/d\mathcal{P}$ ), where  $\sigma$  is the reaction cross section and  $\sigma_L$  is the part of cross section corresponding to mesons with longitudinal polarization (zero helicity state). The limiting values  $\alpha = 1$  and  $\alpha = -1$  refer to the totally transverse and totally longitudinal polarizations. We will be interested in the behavior of  $\alpha$  as a function of the  $\Upsilon$  transverse momentum:  $\mathcal{P} \equiv |\mathbf{p}_T|$ . The experimental definition of  $\alpha$  is based on measuring the angular distributions of the decay leptons  $d\Gamma(\Upsilon \rightarrow \mu^+ \mu^-)/d\cos\theta \sim 1 + \alpha \cos^2\theta$ , where  $\theta$  is the polar angle of the final state muon measured in the decaying meson rest frame.

The results of our calculations for the kinematic conditions of the Tevatron and LHC are displayed in Fig. 12. In both cases, the integration limits over rapidity were adjusted to the experimental acceptances of CDF ( $|y_\Upsilon| < 0.6$ ) at the Tevatron and ATLAS ( $|y_\Upsilon| < 2.5$ ) at the LHC. The upper panels show the predicted transverse momentum distributions. Separately shown are the contributions from the direct (dashed lines) and  $P$ -wave decay (dotted lines) mechanisms.

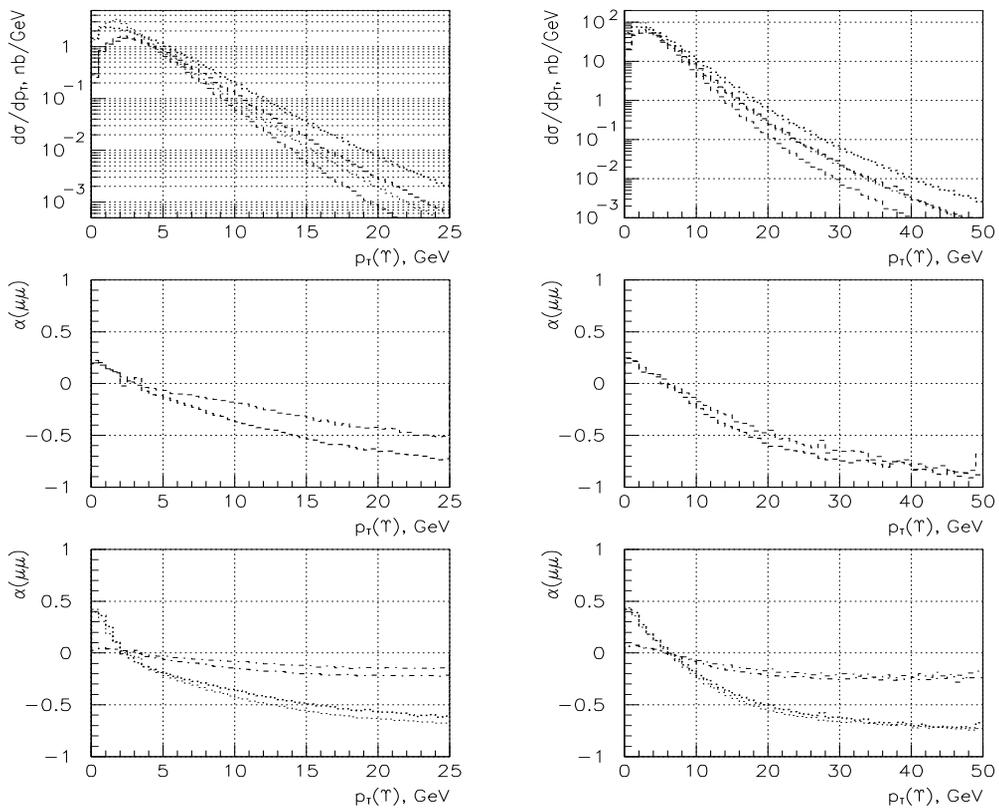


Fig. 12: Predictions on the production of  $\Upsilon$  mesons at the Tevatron (left panel) and LHC (right panel). Thick lines, JB parametrization; thin lines, dGRV parametrization. (a) Transverse momentum distribution. (b) Spin alignment parameter  $\alpha$  for the direct contribution. (c) Spin alignment parameter  $\alpha$  with feed-down from  $\chi_b$  decays taken into account. Dotted lines, the quark spin conservation hypothesis; dash-dotted lines, the full depolarization hypothesis.

As far as the decays of  $P$ -wave states are concerned, nothing is known on the polarisation

properties of these decays. If we assume that the quark spin is conserved in radiative transitions, and the emission of a photon only changes the quark orbital momentum (as it is known to be true in the electric dipole transitions in atomic physics,  $\Delta S = 0$ ,  $\Delta L = \pm 1$ ), then the predictions on  $\alpha$  appear to be similar to those made for the direct channel (lower panels in Fig. 12, dotted curves). If, on the contrary, we assume that the transition  $\chi_b \rightarrow \Upsilon + \gamma$  leads to complete depolarization, then we arrive at a more moderate behavior of the parameter  $\alpha$  (dash-dotted curves in Fig. 12).

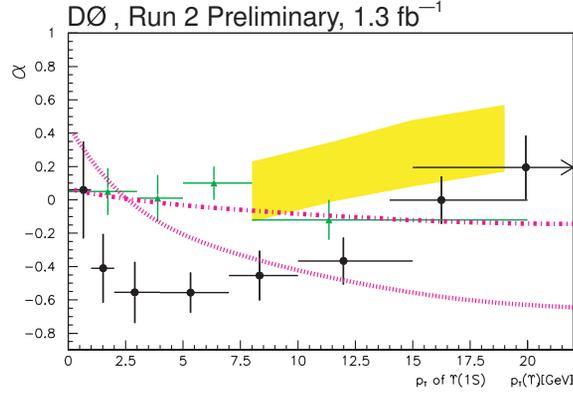


Fig. 13: Spin alignment parameter  $\alpha$  at the Tevatron. Solid curve, quark spin conservation hypothesis; dash-dotted curves, full depolarization hypothesis; yellow band, NRQCD predictions. Green and black points, DØ Run 1 and Run 2 experimental data.

The preliminary results on the  $J/\psi$  polarization at the Tevatron obtained by the collaborations E537 [195] and CDF [196] point to longitudinal polarization with the average value of spin alignment parameter  $\alpha \approx -0.2$  over the whole range of  $J/\psi$  transverse momentum  $p_T$ . In Fig. 13 our results [194] are compared with the preliminary data on the spin alignment of  $\Upsilon$  mesons obtained by the DØ collaboration [197].

A state with purely direct production mechanism in the bottomonium family is the  $\Upsilon(3S)$  meson. The calculations presented here are also valid for this state, except the lower total cross section (by an approximate factor of 1/3) because of the correspondingly lower value of the wave function

At the LHC energies, the theoretical predictions possess less sensitivity to the choice of unintegrated gluon distributions. The purest probe is provided by the polarization of  $\Upsilon(3S)$  mesons.

## 9 $B_c$ and double heavy baryon production and decays

Author: A. Likhoded

$B_c$ -meson is the heaviest of the stable under strong interaction mesons. Because of its unique properties the study of its production and decay processes can be used to check current models of quark dynamics.

There are 16 narrow ( $\bar{b}c$ ) states below the threshold of  $\bar{B}D$ -pair production. In contrast to

$(c\bar{c})$  and  $(b\bar{b})$  systems there are no strong annihilation channel for  $(b\bar{c})$ -mesons, so excited states can decay only to the ground states with the emission of photons and  $\pi$ -mesons.

Experimental value for ground state mass is  $M_{B_c} = 6276.6 \pm 4 \pm 2.7 \text{ MeV}/c^2$  was measured recently by CDF collaboration in exclusive decay  $B_c \rightarrow J/\psi\pi$  [198]. It is in good agreement with theoretical predictions [199] within experimental and theoretical errors. Semileptonic decay mode was used recently by D0 and CDF collaborations to measure  $B_c$ -meson lifetime [200]

$$\tau_{B_c} = 0.448_{-0.096}^{+0.123} \pm 0.121 \text{ ps.}$$

This value is in good agreement with theoretical calculations based on operator product expansion (OPE), potential models and QCD sum rules [201]. This lifetime is caused mainly by decays of  $c$ -quark (70%), while contribution of  $b$ -quark decays and weak annihilation are 20% and 10%, respectively. It should be noted, that observed by D0 and CDF collaboration decay modes are connected with  $b$ -quark decays.

Since both constituent quarks in  $B_c$  are heavy, one can use perturbative QCD for calculation of  $B_c$  production cross section. The only nonperturbative parameter on this cross section, the value of  $B_c$  wave function at the origin, can be obtained using potential models. In this point  $B_c$ -meson production differs dramatically from production of  $B$ - and  $D$ -mesons.

In  $e^+e^-$  annihilation theory predicts usual pattern of  $b$ -quark fragmentation  $b \rightarrow B_c + X$ , with calculable fragmentation functions. In  $\gamma\gamma \rightarrow B_c + X$ ,  $\gamma g \rightarrow B_c + X$  and  $gg \rightarrow B_c + X$  processes, on the other hand, there is strong violation of fragmentation picture for large enough transverse momentum. The factorization formula

$$\frac{d\sigma}{dp_T} = \int \frac{d\hat{\sigma}(\mu, gg \rightarrow b\bar{b})}{dk_T} \Big|_{k_T=p_T/x} D^{b \rightarrow B_c+X}(x, \mu) \frac{dx}{x}$$

is valid only for very large values ( $p_T \gtrsim 40 \text{ GeV}$ ). As a result, to describe experimentally interesting values of  $B_c$  transverse momentum in these processes one needs to calculate total amplitude sets: 20 amplitudes for  $\gamma\gamma$ -, 24 for  $\gamma g$ - and 36 for  $gg$ -subprocesses [202].

A rough estimate of total contribution to  $B_c$  production cross section (including feed-down from excited states) gives the value of order  $10^{-3}$  of the cross section of  $B$ -meson production. CDF and D0 collaborations give their results on  $B_c$  production cross section ( $\sigma_{B_c}$ ) in the form of the ratio over the cross section of  $B$ -meson production ( $\sigma_B$ ):

$$R_e = \frac{\sigma_{B_c} \text{Br}(B_c \rightarrow J/\psi e^+ \nu_e)}{\sigma_B \text{Br}(B_c \rightarrow J/\psi K^\pm)} = 0.282 \pm 0.0038 \pm 0.074$$

in the kinematical region  $m_T(B) > 4.0 \text{ GeV}$  and  $|y(B)| < 1.0$ . Similar result for  $B_c \rightarrow J/\psi \mu^\pm \nu_\mu$  decay is

$$R_\mu = 0.249 \pm 0.045_{-0.076}^{+0.107}.$$

We believe that these results contradict theoretical estimates. Using known branching fractions  $Br(B \rightarrow J/\psi K^\pm) \simeq 1 \cdot 10^{-3}$  and  $Br(B_c \rightarrow J/\psi e^\pm \nu_e) \simeq 2 \cdot 10^{-3}$  one can see, that

Mode	BR, %	Mode	BR, %	Mode	BR, %
$B_c^+ \rightarrow \eta_c e^+ \nu$	0.75	$B_c^+ \rightarrow J/\psi K^+$	0.011	$B_c^+ \rightarrow B_s^0 K^+$	1.06
$B_c^+ \rightarrow \eta_c \tau^+ \nu$	0.23	$B_c \rightarrow J/\psi K^{*+}$	0.022	$B_c^+ \rightarrow B_s^{*0} K^+$	0.37
$B_c^+ \rightarrow \eta_c' e^+ \nu$	0.041	$B_c^+ \rightarrow D^+ \bar{D}^0$	0.0053	$B_c^+ \rightarrow B_s^0 K^{*+}$	–
$B_c^+ \rightarrow \eta_c' \tau^+ \nu$	0.0034	$B_c^+ \rightarrow D^+ \bar{D}^{*0}$	0.0075	$B_c^+ \rightarrow B_s^{*0} K^{*+}$	–
$B_c^+ \rightarrow J/\psi e^+ \nu$	1.9	$B_c^+ \rightarrow D^{*+} \bar{D}^0$	0.0049	$B_c^+ \rightarrow B^0 \pi^+$	1.06
$B_c^+ \rightarrow J/\psi \tau^+ \nu$	0.48	$B_c^+ \rightarrow D^{*+} \bar{D}^{*0}$	0.033	$B_c^+ \rightarrow B^0 \rho^+$	0.96
$B_c^+ \rightarrow \psi' e^+ \nu$	0.132	$B_c^+ \rightarrow D_s^+ \bar{D}^0$	0.00048	$B_c^+ \rightarrow B^{*0} \pi^+$	0.95
$B_c^+ \rightarrow \psi' \tau^+ \nu$	0.011	$B_c^+ \rightarrow D_s^{*+} \bar{D}^{*0}$	0.00071	$B_c^+ \rightarrow B^{*0} \rho^+$	2.57
$B_c^+ \rightarrow D^0 e^+ \nu$	0.004	$B_c^+ \rightarrow D_s^{*+} \bar{D}^0$	0.00045	$B_c^+ \rightarrow B^0 K^+$	0.07
$B_c^+ \rightarrow D^0 \tau^+ \nu$	0.002	$B_c^+ \rightarrow D_s^{*+} \bar{D}^{*0}$	0.0026	$B_c^+ \rightarrow B^0 K^{*+}$	0.015
$B_c^+ \rightarrow D^{*0} e^+ \nu$	0.018	$B_c^+ \rightarrow \eta_c D_s^+$	0.86	$B_c^+ \rightarrow B^{*0} K^+$	0.055
$B_c^+ \rightarrow D^{*0} \tau^+ \nu$	0.008	$B_c^+ \rightarrow \eta_c D_s^{*+}$	0.26	$B_c^+ \rightarrow B^{*0} K^{*+}$	0.058
$B_c^+ \rightarrow B_s^0 e^+ \nu$	4.03	$B_c^+ \rightarrow J/\psi D_s^+$	0.17	$B_c^+ \rightarrow B^+ \bar{K}^0$	1.98
$B_c^+ \rightarrow B_s^{*0} e^+ \nu$	5.06	$B_c^+ \rightarrow J/\psi D_s^{*+}$	1.97	$B_c^+ \rightarrow B^+ \bar{K}^{*0}$	0.43
$B_c^+ \rightarrow B^0 e^+ \nu$	0.34	$B_c^+ \rightarrow \eta_c D^+$	0.032	$B_c^+ \rightarrow B^{*+} \bar{K}^0$	1.60
$B_c^+ \rightarrow B^{*0} e^+ \nu$	0.58	$B_c^+ \rightarrow \eta_c D^{*+}$	0.010	$B_c^+ \rightarrow B^{*+} \bar{K}^{*0}$	1.67
$B_c^+ \rightarrow \eta_c \pi^+$	0.20	$B_c^+ \rightarrow J/\psi D^+$	0.009	$B_c^+ \rightarrow B^+ \pi^0$	0.037
$B_c^+ \rightarrow \eta_c \rho^+$	0.42	$B_c^+ \rightarrow J/\psi D^{*+}$	0.074	$B_c^+ \rightarrow B^+ \rho^0$	0.034
$B_c^+ \rightarrow J/\psi \pi^+$	0.13	$B_c^+ \rightarrow B_s^0 \pi^+$	16.4	$B_c^+ \rightarrow B^{*+} \pi^0$	0.033
$B_c^+ \rightarrow J/\psi \rho^+$	0.40	$B_c^+ \rightarrow B_s^0 \rho^+$	7.2	$B_c^+ \rightarrow B^{*+} \rho^0$	0.09
$B_c^+ \rightarrow \eta_c K^+$	0.013	$B_c^+ \rightarrow B_s^{*0} \pi^+$	6.5	$B_c^+ \rightarrow \tau^+ \nu_\tau$	1.6
$B_c^+ \rightarrow \eta_c K^{*+}$	0.020	$B_c^+ \rightarrow B_s^{*0} \rho^+$	20.2	$B_c^+ \rightarrow c \bar{s}$	4.9

 Table 1: Branching fractions of exclusive  $B_c$  decay modes [203]

in this kinematical region the ratio

$$\frac{\sigma(B_c)}{\sigma(B)} = R_e \frac{\text{Br}(B \rightarrow J/\psi K^\pm) \text{Br}(b \rightarrow B^\pm)}{\text{Br}(B_c \rightarrow J/\psi e^\pm \nu_e)} = \frac{0.282 \cdot 10^{-3} \cdot 0.5}{2 \cdot 10^{-2}} = 0.7 \cdot 10^{-2},$$

that is about an order of magnitude higher than theoretical estimates.

Using CTEQ5L gluon distribution functions and perturbative calculation of  $gg \rightarrow B_c + X$ , we obtained about  $0.8 \mu b$  for  $B_c$ -meson production cross section at LHC. It includes contributions from  $1S_0$  ( $0.19 \mu b$ ),  $1S_1$  ( $0.47 \mu b$ ),  $2S_0$  ( $0.05 \mu b$ ) and  $2S_1$  ( $0.11 \mu b$ ) states. After summing over all spin states we can see, that the whole contribution of  $P$ -wave levels is equal to 7% of  $S$ -state cross section.

At LHC with luminosity  $\mathcal{L} = 10^{34} \text{ cm}^2 \text{ s}^{-1}$  and  $\sqrt{s} = 14 \text{ TeV}$  one can expect  $4.5 \cdot 10^{10}$   $B_c^+$  events per year. As it is clear from Table 1, branching fractions of main semileptonic and hadronic decay modes are large enough for reliable observation of  $B_c$  meson.

## 10 Testing time-reversal and CP symmetry with $\Lambda_b$ decays

Author: Z. J. Ajaltouni

## 10.1 Introduction

Time-reversal (TR) is a fundamental symmetry in many branches of Physics, principally nuclear and particle Physics. Testing its validity or, conversely, searching for its violation, is an important task similar to CP symmetry violation. Few years ago, important experimental results showing clear evidence for TR violation in  $K^0 - \bar{K}^0$  oscillations have been claimed both by CP-LEAR and K-TeV experiments [204]. Then, this research has been extended to the  $B$ -meson system by BaBar and Belle collaborations.

Another source of TR violation could be looked for in particular decays of *hyperons*, as suggested by R. Gatto after the discovery of parity violation in  $\beta$  decay [205]. If we replace the  $s$ -quark belonging to an hyperon by a  $b$ -quark, analogous tests can be performed with *beauty baryons*, like  $\Lambda_b$ ,  $\Sigma_b$ , etc. With the advent of the LHC, it is expected that 10% of the  $b\bar{b}$  pairs produced in proton-proton collisions at  $\sqrt{s} = 14$  TeV will hadronize into beauty baryons  $\mathcal{B}_b$ , and approximately 90% of the  $\mathcal{B}_b$  will be dominated by  $\Lambda_b$  or  $\bar{\Lambda}_b$ . In the framework of the LHCb experiment whose average luminosity will be  $\mathcal{L} = 2 \times 10^{32} \text{cm}^{-2} \text{s}^{-1}$ , roughly  $10^{11}$  beauty baryons will be produced each year.

## 10.2 Features of Time-Reversal

TR operator changes the sign of momentum  $\vec{p}$  and spin  $\vec{s}$  of any particle and leaves its coordinates  $\vec{r}$  invariant. Any triple product  $(\vec{v}_i \times \vec{v}_j) \cdot \vec{v}_k$  with  $\vec{v}_{i,j,k} = \vec{p}$  or  $\vec{s}$  will be *odd* under TR; a non-vanishing value of this observable being a sign of TR violation (TRV). However, an inevitable physical process as strong Final State Interactions (FSI) appears when examining hadronic decays. FSI modify particle wave-functions and generate an additional phase-shift,  $\delta_S$ , to the decay amplitude; the existence of the phase  $\delta_S$  could simulate a  $T$ -odd effect. Being aware of this issue, we developed a phenomenological model describing the decay  $\Lambda_b \rightarrow \Lambda V(1^-)$  and used it in our search for TRV supposing that FSI are negligible. Thus a non-vanishing  $T$ -odd observable will be considered as a serious sign of TRV. In the following, emphasis will be put on TR processes and, because of the delicate problem of CP study in  $\Lambda_b - \bar{\Lambda}_b$  system, only a recent reference will be mentioned [206].

## 10.3 Kinematics and Dynamics of $\Lambda_b \rightarrow \Lambda V(1^-)$ Decays

Different observables can be constructed in order to test TR; the main one being the polarization-vectors of the intermediate resonances coming from  $\Lambda_b$  decays like  $\Lambda(1/2^+)$  and  $V = \rho^0, \omega, J/\psi$ , the vector-meson  $V$  being mainly the  $J/\psi$  decaying into  $\mu^- \mu^+$ . A rigorous study of these decays requires the *helicity formalism* of Jacob-Wick-Jackson which includes the  $\Lambda_b$  initial polarization expressed by its polarization density-matrix (PDM) [207]. Full calculations permit to deduce the  $\Lambda$  angular distributions in an appropriate  $\Lambda_b$  rest-frame. It is given by:  $\frac{d\sigma}{d\Omega} = 1 + \alpha_{As}^{\Lambda_b} \vec{\mathcal{P}}^{\Lambda_b} \cdot \hat{p}$ , where  $\alpha_{As}^{\Lambda_b}$  is the decay asymmetry parameter of the  $\Lambda_b$  resonance,  $\vec{\mathcal{P}}^{\Lambda_b}$  is its polarization-vector and  $\hat{p}$  is the unit-vector parallel to  $\Lambda$  momentum.

A special dynamical model has been performed in order to compute the decay amplitude [208]. It is divided into two main parts : (i) In the framework of the *factorization hypothesis*, the Operator Product Expansion (OPE) techniques are used in order to evaluate both the *soft* (non-perturbative) contributions and the *hard* (perturbative) ones to the hadronic matrix element; the color number

$N_c$  is left *free*. (ii) The form-factors arising in the matrix element are computed by means of the Heavy Quark Effective Theory (HQET) and corrections of order  $\mathcal{O}(1/m_b)$  are performed. Finally, both tree and penguin diagrams have been taken into account in our model.

#### 10.4 Main Physical Results

• In order to test the model, the branching ratio  $BR(\Lambda_b \rightarrow \Lambda J/\psi)$  and other ones are computed according to the effective color number,  $N_c^{eff}$ , and compared to the experimental data.

$N_c^{eff}$	2	2.5	3	3.5
$\Lambda J/\psi$	$8.95 \times 10^{-4}$	$2.79 \times 10^{-4}$	$0.62 \times 10^{-4}$	$0.03 \times 10^{-4}$
$\Lambda \rho^0$	$1.62 \times 10^{-7}$	$1.89 \times 10^{-7}$	$2.2 \times 10^{-7}$	$2.4 \times 10^{-7}$
$\Lambda \omega$	$22.3 \times 10^{-7}$	$4.75 \times 10^{-7}$	$0.2 \times 10^{-7}$	$0.64 \times 10^{-7}$

Table 2: Branching ratio,  $\mathcal{BR}$ , for  $\Lambda_b \rightarrow \Lambda J/\Psi$ ,  $\Lambda_b \rightarrow \Lambda \rho^0$  and  $\Lambda_b \rightarrow \Lambda \omega$ .

The experimental value,  $BR(\Lambda_b \rightarrow \Lambda J/\psi) = (4.7 \pm 2.1 \pm 1.9) \times 10^{-4}$  (PDG 2006), favours the range of values  $2.0 \leq N_c^{eff} \leq 3.0$ .

• Other essential parameters like  $\Lambda_b$  asymmetry,  $\Lambda$  polarization and its non-diagonal matrix element, and the probability of longitudinal polarization for each vector meson can also be obtained :

Parameter	$\Lambda \rho^0 - \omega$	$\Lambda J/\psi$
$\alpha_{AS}^{\Lambda_b}$	0.194	0.490
$\mathcal{P}^\Lambda$	-0.21	-0.17
$\rho_{+-}^\Lambda$	0.31	0.25
$\rho_{00}^V$	0.79	0.66

#### 10.5 Direct Test of Time-Reversal

**Special Angles :**

We define  $\vec{n}_\Lambda$  and  $\vec{n}_V$  respectively as the unit normal vectors to  $\Lambda$  and  $V$  decay planes in the  $\Lambda_b$  rest-frame,  $\vec{e}_Z$  being the quantization axis.

$$\vec{n}_\Lambda = \frac{\vec{p}_p \times \vec{p}_\pi}{|\vec{p}_p \times \vec{p}_\pi|}, \quad \vec{n}_V = \frac{\vec{p}_{l^+} \times \vec{p}_{l^-}}{|\vec{p}_{l^+} \times \vec{p}_{l^-}|}, \quad \text{or} \quad \vec{n}_V = \frac{\vec{p}_{h^+} \times \vec{p}_{h^-}}{|\vec{p}_{h^+} \times \vec{p}_{h^-}|}$$

Those vectors are **even** under TR. But the cosine and the sine of their azimuthal angles defined by :

$$\vec{u}_i = \frac{\vec{e}_Z \times \vec{n}_i}{|\vec{e}_Z \times \vec{n}_i|}, \quad \cos \phi_{(n_i)} = \vec{e}_Y \cdot \vec{u}_i, \quad \sin \phi_{(n_i)} = \vec{e}_Z \cdot (\vec{e}_Y \times \vec{u}_i),$$

with  $\phi_{(n_i)} = \phi_{\vec{n}_\Lambda}$ ,  $\phi_{\vec{n}_V}$  are both **odd** under TR. Their distributions exhibit asymmetries which depend directly on the  $\Lambda$  azimuthal angle distribution whose analytical expression is given by:

$$d\sigma/d\phi \propto 1 + \frac{\pi}{2} \alpha_{As}^\Lambda \left( \Re e(\rho_{+-}^{\Lambda_b}) \cos \phi - \Im m(\rho_{+-}^{\Lambda_b}) \sin \phi \right).$$

The initial  $\Lambda_b$  PDM being unknown, we make the following hypothesis in our simulations :  $\mathcal{P}^{\Lambda_b} = 100\%$  and  $\Re e(\rho_{+-}^{\Lambda_b}) = -\Im m(\rho_{+-}^{\Lambda_b}) = \sqrt{2}/2$ . The following asymmetries are obtained [206] :

Asymmetries	$\Lambda\rho^0 - \omega$	$\Lambda J/\psi$
$AS(\cos \phi_{\vec{n}_\Lambda})$	$(2.4 \pm 0.3)\%$	$(5.2 \pm 0.3)\%$
$AS(\sin \phi_{\vec{n}_\Lambda})$	$-(2.7 \pm 0.3)\%$	$-(5.0 \pm 0.3)\%$

### Vector-Polarizations

In a second step, vector-polarizations have been carefully examined, mainly by considering a new frame related to each resonance  $R_i$  and defined as follows:

$$\vec{e}_L = \frac{\vec{p}}{p}, \quad \vec{e}_T = \frac{\vec{e}_Z \times \vec{e}_L}{|\vec{e}_Z \times \vec{e}_L|}, \quad \vec{e}_N = \vec{e}_T \times \vec{e}_L.$$

Each vector-polarization  $\vec{\mathcal{P}}^{(i)}$  can be expanded on the new basis by writing:  $\vec{\mathcal{P}}^{(i)} = P_L^{(i)} \vec{e}_L + P_N^{(i)} \vec{e}_N + P_T^{(i)} \vec{e}_T$ , with  $P_j^{(i)} = \vec{\mathcal{P}}^{(i)} \cdot \vec{e}_j$  and  $j = L, N, T$ . These components as well as the basis vectors  $\vec{e}_L, \vec{e}_T$  and  $\vec{e}_N$ , are studied under parity and time-reversal operations. The results are straightforward:  $P_L$  and  $P_T$  are both *Parity*-odd and *T*-even, while  $P_N$  is *Parity*-even but *T*-**odd**.

So, if the normal component  $P_N$  is not equal to zero, it would be a signal of TR violation.

### 10.6 Conclusion

The process  $\Lambda_b \rightarrow \Lambda J/\psi$  is a promising channel to look for the validity of TR symmetry at LHC energies. Complete kinematical calculations have been performed by stressing the importance of the resonance polarizations. Our dynamics model is very realistic, because it is based on the OPE formalism and completed by HQET for the computation of the form-factors. An extension of these calculations is under study in order to perform rigorous tests of both CP and TR symmetries among beauty baryons in a model-independent way [209].

## 11 Production and detection of massive exotic hadrons

*Authors: D. Milstead and O. Piskounova*

Exotic stable massive particles are proposed in many models of physics beyond the Standard Model. Understanding their interactions in matter is critical for any search. This paper outlines a model for the scattering of stable massive hadrons which is based on Regge phenomenology and the quark gluon string model.

## 11.1 Introduction

Searches for exotic stable<sup>15</sup> massive particles (SMPs) are performed at colliders as a matter of routine whenever a new collision energy is reached [210]. An additional motivation to make such searches at the LHC arises from the hierarchy problem, proposed solutions to which suggest that new physics processes may be manifest at TeV energies; indeed SMPs are predicted in a number of exotic physics models, such as supersymmetry [210]. Prior to data taking it is important to establish that LHC experiments are able both to detect and extract the quantum numbers of any SMP which may be observed. To do this, an understanding of the interactions of SMPs in matter is needed. As part of this workshop a model [211] has been developed for the scattering of hadronic SMPs (termed  $H$ -hadrons) which uses Regge phenomenology [212] and the quark gluon string model (QGSM) [213]. This work has clear implications for future searches using HERA data and the interpretation of earlier searches.

## 11.2 Interactions of $H$ -Hadrons in Matter

A qualitative picture of the scattering process can be built up [214]. The heavy exotic quark will be a spectator, and the low energy light quark system is involved in the interaction. Regge phenomenology and the QGSM are thus appropriate tools with which the interactions of exotic hadrons in matter can be explored. Fig. 14 shows the predicted cross section for the interaction of a  $H$ -meson with a stationary nucleon in a nucleus comprising equal amounts of protons and neutrons as a function of the Lorentz factor  $\gamma$  of the  $H$ -meson. Reggeon and pomeron contributions are shown separately.

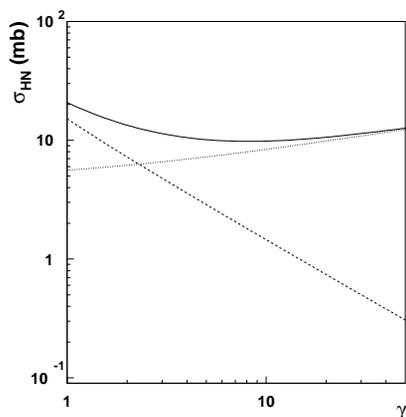


Fig. 14: Pomeron (dotted) and reggeon (dashed) contributions to the exotic-meson-nucleon cross section. The sum of the two processes is shown as a solid line.

Exotic hadrons which contain a light constituent anti-quark, eg  $H_{Q\bar{q}}$  or a  $H_{\bar{Q}q\bar{q}}$  can undergo pomeron and reggeon exchanges. Conversely, hadrons containing a light constituent quark ( $H_{\bar{Q}q}$ ,  $H_{Qqq}$ ) can only undergo pomeron exchange. Anti-baryons and baryons may undergo both

<sup>15</sup>The term stable implies a particle will not decay as it traverses a detector.

reggeon and pomeron exchange, and pomeron exchange only processes, respectively. The overall cross sections for interactions involving baryons and anti-baryons is estimated by doubling the pomeron contribution to the meson cross sections shown in Fig. 14 to take into account the extra light quark contribution. The reggeon contribution to anti-baryon interactions is set to twice the value for meson scattering together with an additional contribution from processes in which exotic anti-baryons can annihilate to exotic mesons and ordinary mesons. This latter contribution is suppressed.

### 11.3 Energy Loss

The PYTHIA [215] program was used to produce samples of stable fourth generation quark pair production events. For reasons of detector acceptance, the  $\beta$  value of the  $H$ -hadrons was restricted to be greater than 0.7 and the pseudorapidity to  $|\eta| < 2.5$  [216]. Using a Monte Carlo method, the  $H$ -hadrons were transported through iron corresponding to the material distribution of the ATLAS detector sub-systems enclosed within the muon detector system. Using a triple regge ansatz [211]  $H$ -hadron energy loss can be estimated. Fig. 15 shows the total energy loss of  $H$ -hadrons after they pass through the detector material. Distributions are presented for  $H$ -hadrons formed from different types of exotic quarks and anti-quarks with masses 200 and 1000 GeV. The distributions are normalised to the total number  $N$  of a given type of  $H$ -hadron satisfying the  $\beta$  and  $\eta$  requirements. There is little difference between them, with a peak around 5 GeV.  $H$ -hadrons containing up-like quarks typically lose more energy than those with down-like quarks owing to the greater fraction of neutral  $H$ -hadrons with down-like quarks.

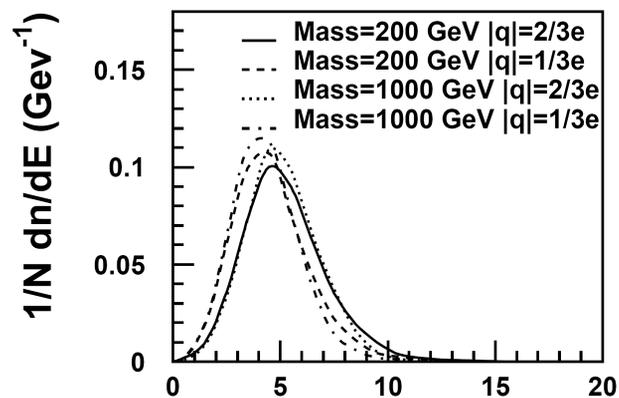


Fig. 15: Total energy loss for  $H$ -hadrons of different types and masses.

### 11.4 Acknowledgements

D. Milstead is a Royal Swedish Academy Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

## References

- [1] W. K. Tung *et al.*, *JHEP* **02**, 053 (2007). hep-ph/0611254.
- [2] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Phys. Lett.* **B652**, 292 (2007). 0706.0459.
- [3] P. M. Nadolsky *et al.*, *Phys. Rev.* **D78**, 013004 (2008). 0802.0007.
- [4] G. Watt, A. D. Martin, W. J. Stirling, and R. S. Thorne, *Recent Progress in Global PDF Analysis*. Preprint 0806.4890, 2008.
- [5] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt (2009). 0901.0002.
- [6] J. C. Collins and W.-K. Tung, *Nucl. Phys.* **B278**, 934 (1986).
- [7] J. C. Collins, *Phys. Rev.* **D58**, 094002 (1998). hep-ph/9806259.
- [8] E. L. M. P. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, *Nucl. Phys.* **B392**, 162 (1993).
- [9] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, *Phys. Rev.* **D50**, 3102 (1994). hep-ph/9312319.
- [10] J. C. Collins, F. Wilczek, and A. Zee, *Phys. Rev.* **D18**, 242 (1978).
- [11] M. Gluck, E. Reya, and A. Vogt, *Eur. Phys. J.* **C5**, 461 (1998). hep-ph/9806404.
- [12] M. Buza, Y. Matiounine, J. Smith, R. Migneron, and W. L. van Neerven, *Nucl. Phys.* **B472**, 611 (1996). hep-ph/9601302.
- [13] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, *Eur. Phys. J.* **C1**, 301 (1998). hep-ph/9612398.
- [14] R. S. Thorne and R. G. Roberts, *Phys. Lett.* **B421**, 303 (1998). hep-ph/9711223.
- [15] R. S. Thorne and R. G. Roberts, *Phys. Rev.* **D57**, 6871 (1998). hep-ph/9709442.
- [16] M. Dittmar *et al.*, *Parton distributions: Summary report for the HERA - LHC workshop*. Preprint hep-ph/0511119, 2005.
- [17] R. M. Barnett, *Phys. Rev. Lett.* **36**, 1163 (1976).
- [18] W. K. Tung, S. Kretzer, and C. Schmidt, *J. Phys.* **G28**, 983 (2002). hep-ph/0110247.
- [19] S. Kretzer, H. L. Lai, F. I. Olness, and W. K. Tung, *Phys. Rev.* **D69**, 114005 (2004). hep-ph/0307022.
- [20] R. S. Thorne, *Phys. Rev.* **D73**, 054019 (2006). hep-ph/0601245.
- [21] A. Chuvakin, J. Smith, and W. L. van Neerven, *Phys. Rev.* **D61**, 096004 (2000). hep-ph/9910250.

- [22] M. Kramer, F. I. Olness, and D. E. Soper, Phys. Rev. **D62**, 096007 (2000).  
hep-ph/0003035.
- [23] S. Kretzer and I. Schienbein, Phys. Rev. **D58**, 094035 (1998). hep-ph/9805233.
- [24] E. L. M. P. Laenen and S. O. Moch, Phys. Rev. **D59**, 034027 (1999).  
hep-ph/9809550.
- [25] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).
- [26] I. Bierenbaum, J. Blümlein, and S. Klein, Acta Phys. Pol. **B39** (2008) 1531 (2008).  
0806.0451.
- [27] M. Buza and W. L. van Neerven, Nucl. Phys. **B500**, 301 (1997). hep-ph/9702242.
- [28] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, *Parton distributions for the LHC*.  
Preprint in preparation, 2008.
- [29] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Phys. Lett.  
**B531**, 216 (2002). hep-ph/0201127.
- [30] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Phys. Lett.  
**B604**, 61 (2004). hep-ph/0410230.
- [31] H. collaboration:, *Measurement of  $F_2^{c\bar{c}}$  and  $F_2^{b\bar{b}}$  using the H1 vertex detector at HERA*.  
34th International Conference on High Energy Physics Philadelphia, 30th July - 5th  
August, 2008.
- [32] S. Alekhin, Phys. Rev. **D68**, 014002 (2003). hep-ph/0211096.
- [33] A. D. Martin, W. J. Stirling, and R. S. Thorne, Phys. Lett. **B636**, 259 (2006).  
hep-ph/0603143.
- [34] B. W. Harris and J. Smith, Nucl. Phys. **B452**, 109 (1995). hep-ph/9503484.
- [35] P. Jimenez-Delgado and E. Reya, *Dynamical NNLO parton distributions*. Preprint  
0810.4274, 2008.
- [36] J. M. Campbell, *Overview of the theory of  $W/Z + jets$  and heavy flavor*. Preprint  
0808.3517, 2008.
- [37] B. A. Kniehl, *Inclusive production of heavy-flavored hadrons at NLO in the GM-VFNS*.  
Preprint 0807.2215, 2008.
- [38] C. Buttar *et al.* (2006). hep-ph/0604120.
- [39] S. J. Brodsky, P. Hoyer, C. Peterson, and N. Sakai, Phys. Lett. **B93**, 451 (1980).
- [40] J. Pumplin, H. L. Lai, and W. K. Tung, Phys. Rev. **D75**, 054029 (2007).  
hep-ph/0701220.

- [41] European Muon Collaboration, J. J. Aubert *et al.*, Nucl. Phys. **B213**, 31 (1983).
- [42] A. D. Martin, R. G. Roberts, W. J. Stirling, and R. S. Thorne, Eur. Phys. J. **C4**, 463 (1998). hep-ph/9803445.
- [43] R. S. Thorne, J. Phys. **G25**, 1307 (1999). hep-ph/9902299.
- [44] OPAL Collaboration, G. Alexander *et al.*, Z. Phys. **C72**, 1 (1996);  
OPAL Collaboration, K. Ackerstaff *et al.*, Eur. Phys. J. **C1**, 439 (1998);  
ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. **C16**, 597 (2000);  
CLEO Collaboration, M. Artuso *et al.*, Phys. Rev. **D70**, 112001 (2004);  
Belle Collaboration, R. Seuster *et al.*, Phys. Rev. **D73**, 032002 (2006).
- [45] G. Kramer and H. Spiesberger, Eur. Phys. J. **C22**, 289 (2001);  
G. Kramer and H. Spiesberger, Eur. Phys. J. **C28**, 495 (2003);  
G. Kramer and H. Spiesberger, Eur. Phys. J. **C38**, 309 (2004).
- [46] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Phys. Rev. **D71**, 014018 (2005);  
B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Eur. Phys. J. **C41**, 199 (2005).
- [47] B. A. Kniehl, G. Kramer, I. Schienbein, and H. Spiesberger, Phys. Rev. Lett. **96**, 012001 (2006).
- [48] T. Kneesch, B. A. Kniehl, G. Kramer, and I. Schienbein, Nucl. Phys. **B799**, 34 (2008).
- [49] M. G. Bowler, Z. Phys. **C11**, 169 (1981).
- [50] SLD Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **84**, 4300 (2000);  
SLD Collaboration, K. Abe *et al.*, Phys. Rev. **D65**, 092006 (2002);  
SLD Collaboration, K. Abe *et al.*, Phys. Rev. **D66**, 079905(E) (2002);  
ALEPH Collaboration, A. Heister *et al.*, Phys. Lett. **B512**, 30 (2001);  
OPAL Collaboration, G. Abbiendi *et al.*, Eur. Phys. J. **C29**, 463 (2003).
- [51] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. **D71**, 021001 (2005);  
CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. **D75**, 012010 (2007).
- [52] J. Binnewies, B. A. Kniehl, and G. Kramer, Phys. Rev. **D58**, 034016 (1998).
- [53] V. G. Kartvelishvili and A. K. Likhoded, Sov. J. Nucl. Phys. **42**, 823 (1985). [Yad. Fiz. **42**, 1306 (1985)].
- [54] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **91**, 241804 (2003).
- [55] U. Aglietti, G. Corcella, and G. Ferrera, Nucl. Phys. **B775**, 162 (2007).
- [56] G. Corcella and G. Ferrera, J. High Energy Phys. **12**, 029 (2007).
- [57] U. Aglietti, G. Ferrera, and G. Ricciardi, Nucl. Phys. **B768**, 85 (2007).

- [58] U. Aglietti, F. Di Lodovico, G. Ferrera, and G. Ricciardi, arXiv:0711.0860. 0711.0860.
- [59] B. Mele and P. Nason, Nucl. Phys. **B617**, 626 (1991).
- [60] M. Cacciari and S. Catani, Nucl. Phys. **B617**, 253 (2001).
- [61] G. Sterman, Nucl. Phys. **B281**, 310 (1987).
- [62] S. Catani and L. Trentadue, Nucl. Phys. **B327**, 323 (1989).
- [63] U. Aglietti and G. Ricciardi, Phys. Rev. **D70**, 114008 (2004).
- [64] D. Shirkov, Nucl. Phys. Proc. Suppl. **152**, 51 (2006).
- [65] ALEPH Collaboration, A. Heister *et al.*, Phys. Lett. **B512**, 30 (2001).
- [66] OPAL Collaboration, G. Abbiendi *et al.*, Eur. Phys. J. **C29**, 463 (2003).
- [67] SLD Collaboration, K. Abe *et al.*, Phys. Rev. Lett. **84**, 4300 (2000).
- [68] ALEPH Collaboration, R. Barate *et al.*, Eur. Phys. J. **C16**, 597 (2000).
- [69] ZEUS Collaboration, S. Chekanov *et al.*, Nucl. Phys. **B700**, 3 (2004). hep-ex/0405065.
- [70] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D60**, 072003 (1999). hep-ex/9903011.
- [71] A. Banfi, G. P. Salam, and G. Zanderighi, Eur. Phys. J. **C47**, 113 (2006). hep-ph/0601139.
- [72] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. **B269**, 432 (1991).
- [73] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).
- [74] S. D. Ellis and D. E. Soper, Phys. Rev. **D48**, 3160 (1993). hep-ph/9305266.
- [75] A. Banfi, G. P. Salam, and G. Zanderighi, JHEP **08**, 062 (2004). hep-ph/0407287.
- [76] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, JHEP **11**, 063 (2001). hep-ph/0109231.
- [77] M. L. Mangano, M. Moretti, and R. Pittau, Nucl. Phys. **B632**, 343 (2002). hep-ph/0108069.
- [78] *Measurement of the inclusive b-jet cross section in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96\text{TeV}$*  (unpublished). Note 8418.
- [79] J. M. Campbell and R. K. Ellis, Phys. Rev. **D62**, 114012 (2000). hep-ph/0006304.

- [80] M. L. Mangano, P. Nason, and G. Ridolfi, Nucl. Phys. **B373**, 295 (1992).
- [81] S. Frixione, P. Nason, and B. R. Webber, JHEP **08**, 007 (2003). hep-ph/0305252.
- [82] G. Corcella *et al.*, JHEP **01**, 010 (2001). hep-ph/0011363.
- [83] A. Banfi, G. P. Salam, and G. Zanderighi, JHEP **07**, 026 (2007). 0704.2999.
- [84] W. T. Giele, E. W. N. Glover, and D. A. Kosower, Nucl. Phys. **B403**, 633 (1993). hep-ph/9302225.
- [85] Z. Nagy, Phys. Rev. Lett. **88**, 122003 (2002). hep-ph/0110315.
- [86] J. Pumplin, H. L. Lai, and W. K. Tung, Phys. Rev. **D75**, 054029 (2007). hep-ph/0701220.
- [87] CDF Collaboration, D. E. Acosta *et al.*, Phys. Rev. **D71**, 092001 (2005). hep-ex/0412006.
- [88] C. Weiser, *A combined secondary vertex based b-tagging algorithm in CMS* (unpublished). CMS NOTE-2006/014.
- [89] M. Sapinski, *Expected performance of ATLAS for measurements of jets, b-jets, tau-jets, and  $E_{T,mis}$*  (unpublished). SN-ATLAS-2002-012, ATL-COM-CONF-2001-006.
- [90] J. Bastos, physics/0702041 (2007). physics/0702041.
- [91] W. Bernreuther, J. Phys. **G35**, 083001 (2008). 0805.1333.
- [92] P. M. Nadolsky *et al.*, Phys. Rev. **D78**, 013004 (2008). 0802.0007.
- [93] P. Nason, S. Dawson, and R. K. Ellis, Nucl. Phys. **B303**, 607 (1988).
- [94] S. Moch and P. Uwer, Phys. Rev. **D78**, 034003 (2008). 0804.1476.
- [95] M. Cacciari, S. Frixione, M. M. Mangano, P. Nason, and G. Ridolfi (2008). 0804.2800.
- [96] N. Kidonakis and R. Vogt, arXiv:0805.3844 (2008). 0805.3844.
- [97] M. Czakon, A. Mitov, and S. Moch, Phys. Lett. **B651**, 147 (2007). 0705.1975.
- [98] M. Czakon, A. Mitov, and S. Moch, Nucl. Phys. **B798**, 210 (2008). 0707.4139.
- [99] M. Czakon, Phys. Lett. **B664**, 307 (2008). 0803.1400.
- [100] J. G. Körner, Z. Merebashvili, and M. Rogal, Phys. Rev. **D73**, 034030 (2006). hep-ph/0511264.
- [101] J. G. Körner, Z. Merebashvili, and M. Rogal, Phys. Rev. **D77**, 094011 (2008). 0802.0106.

- [102] S. Dittmaier, P. Uwer, and S. Weinzierl, *Phys. Rev. Lett.* **98**, 262002 (2007).  
hep-ph/0703120.
- [103] M. Caffo, H. Czyz, S. Laporta, and E. Remiddi, *Nuovo Cim.* **A111**, 365 (1998).  
hep-th/9805118.
- [104] E. Laenen, S. Riemersma, J. Smith, and W. L. van Neerven, *Nucl. Phys.*  
**B392**, 162 (1993);  
S. Riemersma, J. Smith, and W. L. van Neerven, *Phys. Lett.* **B347**, 143 (1995).  
hep-ph/9411431.
- [105] S. Alekhin and J. Blümlein, *Phys. Lett.* **B594**, 299 (2004). hep-ph/0404034.
- [106] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, *Nucl. Phys.*  
**B485**, 420 (1997). hep-ph/9608342.
- [107] I. Bierenbaum, J. Blümlein, and S. Klein, in preparation.
- [108] S. Alekhin, K. Melnikov, and F. Petriello, *Phys. Rev.* **D74**, 054033 (2006).  
hep-ph/0606237.
- [109] Blümlein, H. J., Böttcher, and A. Guffanti, *Nucl. Phys.* **B774**, 182 (2007).  
hep-ph/0607200.
- [110] M. Buza, Y. Matiounine, J. Smith, R. Migneron, and W. L. van Neerven, *Nucl. Phys.*  
**B472**, 611 (1996). hep-ph/9601302.
- [111] J. A. M. Vermaseren, A. Vogt, and S. Moch, *Nucl. Phys.* **B724**, 3 (2005).  
hep-ph/0504242.
- [112] I. Bierenbaum, J. Blümlein, and S. Klein, *Nucl. Phys.* **B780**, 40 (2007).  
hep-ph/0703285.
- [113] I. Bierenbaum, J. Blümlein, and S. Klein, *Phys. Lett.* **B648**, 195 (2007).  
hep-ph/0702265.
- [114] J. Blümlein, A. De Freitas, W. L. van Neerven, and S. Klein, *Nucl. Phys.*  
**B755**, 272 (2006). hep-ph/0608024.
- [115] I. Bierenbaum, Blümlein, S. J., Klein, and C. Schneider, *Nucl. Phys.* **B803**, 1 (2008).  
0803.0273.
- [116] S. A. Larin, T. van Ritbergen, and J. A. M. Vermaseren, *Nucl. Phys.* **B427**, 41 (1994);  
S. A. Larin, P. Nogueira, T. van Ritbergen, and J. A. M. Vermaseren, *Nucl. Phys.*  
**B492**, 338 (1997). hep-ph/9605317;  
A. Retey and J. A. M. Vermaseren, *Nucl. Phys.* **B604**, 281 (2001). hep-ph/0007294;  
J. Blümlein and J. A. M. Vermaseren, *Phys. Lett.* **B606**, 130 (2005).  
hep-ph/0411111;  
S. Moch, J. A. M. Vermaseren, and A. Vogt, *Nucl. Phys.* **B688**, 101 (2004).  
hep-ph/0403192.

- [117] A. Vogt, S. Moch, and J. A. M. Vermaseren, Nucl. Phys. **B691**, 129 (2004).  
hep-ph/0404111.
- [118] J. A. Gracey, Phys. Lett. **B322**, 141 (1994). hep-ph/9401214.
- [119] N. Gray, D. J. Broadhurst, W. Grafe, and K. Schilcher, Z. Phys. **C48**, 673 (1990).
- [120] M. Buza, Y. Matiounine, J. Smith, and W. L. van Neerven, Eur. Phys. J. **C1**, 301 (1998).  
hep-ph/9612398.
- [121] I. Bierenbaum, J. Blümlein, and S. Klein, arxiv:0806.4613 [hep-ph] (2008).  
0806.4613.
- [122] B. A. Ovrut and H. J. Schnitzer, Nucl. Phys. **B179**, 381 (1981);  
B. A. Ovrut and H. J. Schnitzer, Nucl. Phys. **B189**, 509 (1981);  
W. Bernreuther and W. Wetzel, Nucl. Phys. **B197**, 228 (1982);  
W. Bernreuther, Ann. Phys. **151**, 127 (1983).
- [123] I. Bierenbaum, J. Blümlein, and S. Klein, Acta Phys. Polon. **B38**, 3543 (2007).  
0710.3348.
- [124] J. Blümlein and S. Kurth, Phys. Rev. **D60**, 014018 (1999). hep-ph/9810241.
- [125] J. A. M. Vermaseren, Int. J. Mod. Phys. **A14**, 2037 (1999). hep-ph/9806280.
- [126] J. Blümlein, Comput. Phys. Commun. **159**, 19 (2004). hep-ph/0311046.
- [127] J. Blümlein and S. Klein, arxiv:0706.2426 [hep-ph] (2007). 0706.2426.
- [128] J. Blümlein, arxiv:0807.0700 [math-ph] (2008). 0807.0700.
- [129] J. Blümlein and N. Kochelev, Phys. Lett. **B381**, 296 (1996). hep-ph/9603397.
- [130] J. Blümlein, V. Ravindran, and W. L. van Neerven, Phys. Rev. **D68**, 114004 (2003).  
hep-ph/0304292.
- [131] P. Nogueira, J. Comput. Phys. **105**, 279 (1993).
- [132] T. van Ritbergen, A. N. Schellekens, and J. A. M. Vermaseren, Int. J. Mod. Phys.  
**A14**, 41 (1999). hep-ph/9802376.
- [133] M. Steinhauser, Comput. Phys. Commun. **134**, 335 (2001). hep-ph/0009029.
- [134] J. A. M. Vermaseren, math-ph/0010025 (2000). math-ph/0010025.
- [135] C. Schneider, Ann. Comb., **9** (1) (2005) 75; Proc. ISSAC'05, (2005) pp. 285 (ACM  
Press); Proc. FPSAC'07, (2007) 1; J. Diffr. Equations Appl., **11** (9) (2005) 799; J.  
Algebra Appl., **6** (3) (2007) 415; Sém. Lothar. Combin. **56** (2007) Article B56b and  
Habilitation Thesis, JKU Linz, (2007).
- [136] J. Blümlein, DESY 08-042.

- [137] J. Blümlein and V. Ravindran, Nucl. Phys. **B716**, 128 (2005). hep-ph/0501178.
- [138] J. Blümlein and V. Ravindran, Nucl. Phys. **B749**, 1 (2006). hep-ph/0604019.
- [139] M. Tentyukov and J. A. M. Vermaseren, hep-ph/0702279 (2007). hep-ph/0702279.
- [140] CTEQ Collaboration, R. Brock *et al.*, Rev. Mod. Phys. **67**, 157 (1995).
- [141] V. N. Gribov and L. N. Lipatov, Sov. J. Nucl. Phys. **15**, 438 (1972). [Yad. Fiz. **15**, 781 (1972)];  
Y. L. Dokshitzer, Sov. Phys. JETP **46**, 641 (1977). [Zh. Eksp. Teor. Fiz. **73**, 1216 (1977)];  
G. Altarelli and G. Parisi, Nucl. Phys. **B126**, 298 (1977).
- [142] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **44**, 443 (1976). [Zh. Eksp. Teor. Fiz. **71**, 840 (1976)];  
I. I. Balitsky and L. N. Lipatov, Sov. J. Nucl. Phys. **28**, 822 (1978). [Yad. Fiz. **28**, 1597 (1978)].
- [143] V. S. Fadin and L. N. Lipatov, Nucl. Phys. **B477**, 767 (1996).
- [144] L. N. Lipatov, Nucl. Phys. **B452**, 452 (1995);  
E. N. Antonov, L. N. Lipatov, E. A. Kuraev, and I. O. Cherednikov, Nucl. Phys. **B721**, 111 (2005);  
L. N. Lipatov and M. I. Vyazovsky, Nucl. Phys. **B597**, 399 (2001).
- [145] M. Gell-Mann, M. L. Goldberger, F. E. Low, E. Marx, and F. Zachariasen, Phys. Rev. **133**, 145B (1964).
- [146] V. S. Fadin and V. E. Sherman, JETP Lett. **23**, 548 (1976);  
V. S. Fadin and V. E. Sherman, JETP **45**, 861 (1977).
- [147] G. T. Bodwin, E. Braaten, and G. P. Lepage, Phys. Rev. **D51**, 1125 (1995). Phys. Rev. **D55**, 5853E (1997).
- [148] M. Cacciari and P. Nason, JHEP **09**, 006 (2003).
- [149] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **79**, 572 (1997);  
CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **79**, 578 (1997);  
CDF Collaboration, T. Affolder *et al.*, Phys. Rev. Lett. **85**, 2886 (2000).
- [150] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005).
- [151] G. T. Bodwin, E. Braaten, and J. Lee, Phys. Rev. **D72**, 014004 (2005).
- [152] P. Hägler, R. Kirschner, A. Schäfer, L. Szymanowski, and O. V. Teryaev, Phys. Rev. **D62**, 071502 (2000);  
P. Hägler, R. Kirschner, A. Schäfer, L. Szymanowski, and O. V. Teryaev, Phys. Rev. Lett. **86**, 1446 (2001).
- [153] F. Yuan and K. T. Chao, Phys. Rev. **D63**, 034006 (2001).

- [154] F. Yuan and K. T. Chao, Phys. Rev. Lett. **87**, 022002 (2001).
- [155] B. A. Kniehl, D. V. Vasin, and V. A. Saleev, Phys. Rev. **D73**, 074022 (2006).
- [156] V. A. Saleev and D. V. Vasin, Phys. Rev. **D68**, 114013 (2003);  
V. A. Saleev and D. V. Vasin, Phys. Atom. Nucl. **68**, 94 (2005). [*Yad. Fiz.* **68**, 95 (2005)].
- [157] J. Blümlein, preprint (1995). DESY 95-121.
- [158] H. Jung and G. P. Salam, Eur. Phys. J. **C19**, 351 (2001).
- [159] M. A. Kimber, A. D. Martin, and M. G. Ryskin, Phys. Rev. **D63**, 114027 (2001).
- [160] ZEUS Collaboration, S. Chekanov *et al.*, Eur. Phys. J. **C27**, 173 (2003).
- [161] H1 Collaboration, C. Adloff *et al.*, Eur. Phys. J. **C25**, 41 (2002).
- [162] V. A. Saleev and D. V. Vasin, Phys. Lett. **B548**, 161 (2002);  
V. A. Saleev, Phys. Rev. **D65**, 054041 (2002).
- [163] V. A. Saleev and N. P. Zotov, Mod. Phys. Lett. **A9**, 151 (1994).
- [164] CDF Collaboration, F. Abe *et al.*, Phys. Rev. Lett. **75**, 4358 (1995);  
CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **88**, 161802 (2002).
- [165] CDF Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. **94**, 232001 (2005).
- [166] E. Braaten, S. Fleming, and A. K. Leibovich, Phys. Rev. **D63**, 094006 (2001).
- [167] B. A. Kniehl, V. A. Saleev, and D. V. Vasin, Phys. Rev. **D74**, 014024 (2006).
- [168] V. A. Saleev and D. V. Vasin, Phys. Part. Nucl. **38**, 635 (2007).
- [169] H1 Collaboration, C. Adloff *et al.*, Nucl. Phys. **B545**, 21 (1999).
- [170] ZEUS Collaboration, J. Breitweg *et al.*, Eur. Phys. J. **C6**, 67 (1999). Phys. Lett. **B481**, 213 (2000).
- [171] V. A. Saleev and D. V. Vasin, Phys. Atom. Nucl. **68**, 95 (2005).
- [172] CDF Collaboration, D. Acosta *et al.*, Phys. Rev. Lett. **91**, 241804 (2003).
- [173] J. C. Collins and R. K. Ellis, Nucl. Phys. **B360**, 3 (1991).
- [174] S. Catani, M. Ciafaloni, and F. Hautmann, Nucl. Phys. **B366**, 135 (1991).
- [175] B. A. Kniehl, A. V. Shipilova, and V. A. Saleev, to be published.
- [176] B. A. Kniehl, A. V. Shipilova, and V. A. Saleev, DESY-08-194 **2008**.  
arXiv:0812.3376[hep-ph].
- [177] CDF Collaboration, F. Abe *et al.*, Phys. Rev. **D55**, 2546 (1997).

- [178] P. Hagler *et al.*, Phys. Rev. **D62**, 071502(R) (2000).
- [179] J. Kwiecinski, A. Martin, and S. A., Phys. Rev. **D56**, 3991 (1997).
- [180] C. Peterson *et al.*, Phys. Rev. **D27**, 105 (1983).
- [181] V. A. Saleev and D. V. Vasin. arXiv:0709.0259.
- [182] M. Kramer, Prog. Part. Nucl. Phys. **47**, 141 (2001). hep-ph/0106120.
- [183] Quarkonium Working Group Collaboration, N. Brambilla *et al.* (2004). hep-ph/0412158.
- [184] B. A. Kniehl and L. Zvirner, Nucl. Phys. **B621**, 337 (2002). hep-ph/0112199.
- [185] S. P. Baranov, Phys. Rev. **D66**, 114003 (2002).
- [186] P. Hagler, R. Kirschner, A. Schafer, L. Szymanowski, and O. V. Teryaev, Phys. Rev. **D63**, 077501 (2001). hep-ph/0008316.
- [187] S. P. Baranov, Phys. Lett. **B428**, 377 (1998).
- [188] S. P. Baranov, A. V. Lipatov, and N. P. Zotov (2001). hep-ph/0106229.
- [189] L. V. Gribov, E. M. Levin, and M. G. Ryskin, Phys. Rept. **100**, 1 (1983).
- [190] S. Catani, M. Ciafaloni, and F. Hautmann, Phys. Lett. **B242**, 97 (1990).
- [191] E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Sov. Phys. JETP **45**, 199 (1977).
- [192] J. Blümlein, J. Phys. **G19**, 1623 (1993).
- [193] M. Gluck, E. Reya, and A. Vogt, Eur. Phys. J. **C5**, 461 (1998). hep-ph/9806404.
- [194] S. P. Baranov and N. P. Zotov, JETP Lett. **86**, 435 (2007). 0707.0253.
- [195] C. Akerlof *et al.*, Phys. Rev. **D48**, 5067 (1993).
- [196] CDF Collaboration, A. A. Affolder *et al.*, Phys. Rev. Lett. **85**, 2886 (2000). hep-ex/0004027.
- [197] D0 Collaboration, V. M. Abazov *et al.* (2008). 0804.2799.
- [198] V. Papadimitriou, AIP Conf. Proc. **815**, 157 (2006). [arXiv:hep-ex/0511043]; M. D. Corcoran, others (for the CDF, and D. Coll.), arXiv:hep-ex/0506061.
- [199] S. S. Gershtein, V. V. Kiselev, A. K. Likhoded, and A. V. Tkabladze, Phys. Usp. **38**, 1 (1995). [Usp. Fiz. Nauk. **165** (1995) 3, arXiv:hep-ph/9504319]; V. V. Kiselev, A. K. Likhoded, and A. V. Tkabladze, Phys. Rev. **D51**, 3613 (1995); S. N. Gupta and J. M. Johnson, Phys. Rev. **D53**, 312 (1996); D. Ebert, R. N. Faustov, and V. O. Galkin, Phys. Rev. **D67**, 014027 (2003);

- N. Brambilla, Y. Sumin, and A. Vairo, Phys. Rev. **D65**, 034001 (2002);  
 S. M. Ikhdaïr and R. Sever, Int. J. Mod. Phys. **A20**, 6509 (2005);  
 S. Godfrey, Phys. Rev. **D70**, 054017 (2004).
- [200] CDF Collaboration, A. Abulencia *et al.*, Phys. Rev. Lett. **97**, 242003 (2006);  
 D0 Collaboration, V. M. Abazov *et al.*, Phys. Rev. Lett. **97**, 021802 (2006).
- [201] I. P. Gouzev, V. V. Kiselev, A. K. Likhoded, V. I. Romanovsky, and O. P. Yushchenko,  
 Phys. Atom. Nucl. **67**, 1559 (2004). [Yad. Fiz. **67**, 1581 (2004), arXiv:hep-ph/0211432];  
 V. V. Kiselev, A. E. Kovalsky, and A. K. Likhoded, Nucl. Phys. **B585**, 353 (2000);  
 C. H. Chang, Int. J. Mod. Phys. **A21**, 777 (2006);  
 V. V. Kiselev, arXiv:hep-ph/0308214.
- [202] V. V. Kiselev, A. K. Likhoded, and A. I. Onishchenko, Nucl. Phys. **B569**, 473 (2000);  
 A. V. Berezhnoi, V. V. Kiselev, and A. K. Likhoded, Z. Phys. **A356**, 79 (1996);  
 K. Kolodziej and R. Rückl, Nucl. Instrum. Meth. **A408**, 33 (1998);  
 C. H. Chang, Y. Q. Chen, G. P. Han, and H. T. Jiang, Phys. Lett. **B364**, 78 (1995).
- [203] S. S. Gershtein and A. K. Likhoded, arXiv:0706.0963.
- [204] L. Alvarez-Gaumé *et al.*, arXiv:hep-ph/9812326.
- [205] R. Gatto, Nucl. Phys. **5**, 183 (1958).
- [206] E. Conte, Thèse de Docteur de l'Université Blaise Pascal **PPCF T 0710** (2007).
- [207] Z. J. Ajaltouni *et al.*, Phys. Lett. **B614**, 165 (2005).
- [208] O. Leitner, Z. J. Ajaltouni, and E. Conte, arXiv:hep-ph/0602043;  
 Z. J. Ajaltouni *et al.*, Nucl. Phys. B Proc. Suppl. **174**, 169 (2007).
- [209] D. S. E. *et al.*, PCCF RI 0802 (2008).
- [210] M. Fairbairn *et al.*, Phys. Rept. **438**, 1 (2007). hep-ph/0611040.
- [211] Y. R. de Boer, A. B. Kaidalov, D. A. Milstead, and O. I. Piskounova, J. Phys. **G35**, 075009 (2008). 0710.3930.
- [212] P. D. B. Collins, Phys. Rept. **1**, 103 (1971);  
 P. D. B. Collins. Cambridge 1977, 445p.
- [213] A. B. Kaidalov, Phys. Lett. **B116**, 459 (1982);  
 A. B. Kaidalov and K. A. Ter-Martirosian, Phys. Lett. **B117**, 247 (1982);  
 A. B. Kaidalov and K. A. Ter-Martirosian, Sov. J. Nucl. Phys. **39**, 979 (1984);  
 A. B. Kaidalov and O. I. Piskunova, Z. Phys. **C30**, 145 (1986).
- [214] M. Drees and X. Tata, Phys. Lett. **B252**, 695 (1990);  
 A. C. Kraan, Eur. Phys. J. **C37**, 91 (2004). hep-ex/0404001.
- [215] T. Sjostrand, S. Mrenna, and P. Skands, JHEP **05**, 026 (2006). hep-ph/0603175.

- [216] A. C. Kraan, J. B. Hansen, and P. Nevski, *Eur. Phys. J. C***49**, 623 (2007).  
hep-ex/0511014.