Proceedings of the

First International Workshop on Multiple Partonic Interactions at the LHC

MPI'08

October 27-31, 2008
Perugia, Italy

Editors: Paolo Bartalini[1], Livio Fanò[2]

(1) National Taiwan University
(2) INFN and Università degli Studi di Perugia

Impressum

Proceedings of the First International Workshop on Multiple Partonic Interactions at the LHC (MPI08)
October 27-31, 2008, Perugia, Italy

Conference homepage
`http://www.pg.infn.it/mpi08`

Slides at
`https://agenda.infn.it/conferenceDisplay.py?confId=599`

Editors:
Paolo Bartalini, Livio Fanò
June 2009
DESY-PROC-2009-06
ISBN 978-3-935702-38-6
ISSN 1435-8077

Published by
Verlag Deutsches Elektronen-Synchrotron
Notkestraße 85
22607 Hamburg
Germany

Organizing Committee

Scientific Advisory Committee:

P. Bartalini (National Taiwan University, Taipei, TW)
J. Butterworth (University College London, London, UK)
L. Fanò (Istituto Nazionale di Fisica Nucleare, Perugia, IT)
R. Field (University of Florida, Gainesville, US)
I. Hinchliffe (Lawrence Berkeley National Laboratory, Berkeley, US)
H. Jung (Deutsches Elektronen-Synchrotron, Hamburg, DE)
S. Lami (Istituto Nazionale di Fisica Nucleare, Pisa, IT)
A. Morsch (European Organization for Nuclear Research, Meyrin, CH)
G. Pancheri (Istituto Nazionale di Fisica Nucleare, Frascati, IT)
M. Schmelling (Max-Planck-Institut fur Kernphysik, Heidelberg, DE)
T. Sjostrand (Lunds Universitet, Lund, SE)
Y. Srivastava (Università degli Studi di Perugia, Perugia, IT)
J. Stirling (Institute for Particle Physics Phenomenology, Durham, UK)
M. Strikman (Pennsylvania State University, University Park, US)
D. Treleani (Università degli Studi di Trieste, Trieste, IT)

Local Advisory Committee:

F. Ambroglini (Università degli Studi di Trieste, Trieste, IT)
G.M. Bilei (Istituto Nazionale di Fisica Nucleare, Perugia, IT)
G. Chiocci (Università degli Studi di Perugia, Perugia, IT)
L. Fanò (Istituto Nazionale di Fisica Nucleare, Perugia, IT)
A. Santocchia (Università degli Studi di Perugia, Perugia, IT)

Preface

The objective of this first workshop on Multiple Partonic Interactions (MPI) at the LHC, that can be regarded as a continuation and extension of the dedicated meetings held at DESY in the years 2006 and 2007, is to raise the profile of MPI studies, summarizing the legacy from the older phenomenology at hadronic colliders and favouring further specific contacts between the theory and experimental communities. The MPI are experiencing a growing popularity and are currently widely invoked to account for observations that would not be explained otherwise: the activity of the Underlying Event, the cross sections for multiple heavy flavour production, the survival probability of large rapidity gaps in hard diffraction, etc. At the same time, the implementation of the MPI effects in the Monte Carlo models is quickly proceeding through an increasing level of sophistication and complexity that in perspective achieves deep general implications for the LHC physics. The ultimate ambition of this workshop is to promote the MPI as unification concept between seemingly heterogeneous research lines and to profit of the complete experimental picture in order to constrain their implementation in the models, evaluating the spin offs on the LHC physics program. The workshop is structured in five sections, with the first one dedicated to few selected hot highlights in the High Energy Physics and directly connected to the other ones: Multiple Parton Interactions (in both the soft and the hard regimes), Diffraction, Monte Carlo Generators and Heavy Ions.

Contents

Part I

Hot Topics

Convenors:

Paolo Bartalini (National Taiwan University)
Yogendra Srivastava (University of Perugia)

Standard Model Higgs Searches at the Tevatron

Ralf Bernhard
Physikalisches Institut, Albert-Ludwigs Universität Freiburg

Abstract
The latest searches for the Standard Model Higgs boson at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV with the DØ and the CDF detectors at the Fermilab Tevatron collider are presented. For the first time since the LEP experiments the sensitivity for a Standard Model Higgs boson has been reached at a Higgs boson mass of 170 GeV/c^2.

1 Introduction

In the Standard Model (SM) of particle physics the Higgs mechanism is responsible for breaking electroweak symmetry, thereby giving mass to the W and Z bosons. It predicts the existence of a heavy scalar boson, the Higgs boson, with a mass that can not be predicted by the SM. Direct searches for the Higgs Boson were performed at the LEP experiments in the process $e^+e^- \rightarrow ZH$ with a centre of mass energy of 206.6 GeV. A direct mass limit at $m_H > 114.4$ GeV/c^2 [1] was set at the 95% confidence level (CL)[1]. This limit is slightly below the maximum available kinematic limit due to a small excess observed in the LEP data.

Indirect limits have been placed on the Higgs boson mass by the LEP, SLD and Tevatron experiments from electroweak precision measurements [2]. The main contribution to these indirect constraints from the Tevatron experiments, DØ and CDF, are the measurements of the W Boson and top quark masses [2]. The dependence of the Higgs mass on these measurements is shown in Figure 1 on the left and the Higgs mass dependence on the measured electroweak precision parameters in Figure 1 on the right. The SM fit yields a best value of $m_H = 84^{+34}_{-26}$ GeV/c^2 [3]. The upper limit on the Higgs mass at 95% CL is $m_H < 154$ GeV/c^2. If the direct mass limit is also taken into account this limit is increased to $m_H < 185$ GeV/c^2.

2 Higgs Searches at the Tevatron

The Tevatron experiments CDF [4] and DØ [5] search for direct Higgs boson production in the mass range above the LEP limit using $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. The relevant processes at these energies are associated Higgs production ($qq' \rightarrow WH$, $q\bar{q} \rightarrow ZH$) and gluon fusion ($gg \rightarrow H$). Typical cross-sections are $\sigma \simeq 0.7 - 0.15$ pb for gluon fusion and $\sigma \simeq 0.2 - 0.02$ pb for associated production at Higgs masses in the range $115 - 200$ GeV/c^2.

The Higgs boson predominantly decays into $b\bar{b}$ quark pairs in the low mass range below 135 GeV/c^2. Hence the signal in the $gg \rightarrow H$ channel is overwhelmed by multi-jet background. This makes the process $gg \rightarrow H$ therefore not a viable search channel at low Higgs boson masses. The WH and ZH channels, where the vector boson decays into leptons, have much lower cross-sections but the lepton tag from the decay of the $W \rightarrow \ell\nu$ or $Z \rightarrow \ell\ell$ and selections

[1] All limits given in this paper are at 95% CL

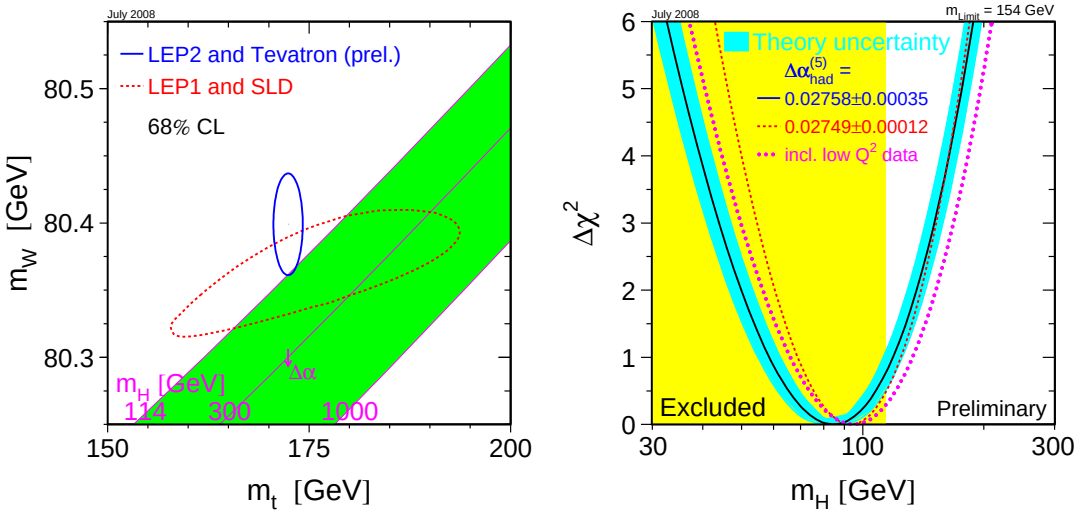

Fig. 1: Constraints on the Higgs mass from precision top and W mass measurements (left) and fit for the Higgs Mass from the W data showing the direct search LEP limit (right)

on missing transverse energy from the neutrino in the decays $W \to \ell\nu$ or $Z \to \nu\nu$ help to reduce the background significantly.

At higher masses, around $m_H = 165$ GeV/c^2, the Higgs boson will predominantly decay into WW pairs. Leptons from the decays of the W bosons and the missing transverse energy are used to reject background, making the channel $gg \to H \to WW$ the most promising search channel in this mass region. A 'hybrid' channel, the associated production with subsequent Higgs decay into (virtual) W pairs, $qq' \to WH \to WWW$, also contributes in the intermediate mass region.

2.1 The Tools

The main tools employed in Higgs searches at the Tevatron are lepton identification and - especially in the low Higgs mass region - jet reconstruction and b jet tagging. The experiments use b jet tagging algorithms that exploit the long lifetime of b hadrons. These algorithms are applied to each jet, searching for tracks with large transverse impact parameters relative to the primary vertex and for secondary vertices formed by tracks in the jet.

To further improve the b jet tagging these variables are used as input to a artificial Neural Network (NN) jet-flavor separator. The NN is trained to separate b quark jets from light flavour jets. By adjusting the minimum requirement on the NN output variable, a range of increasingly stringent b tagging operating points is obtained, each with a different signal efficiency and purity. Using this tool at DØ , b tagging efficiencies have been improved by 33% while keeping the rate of falsely identified light flavor jets (mistags) low. The efficiencies range between 40-70% for b

jets at a low mistag rates between 0.5-3% for light flavor jets.

Almost all Higgs searches at the Tevatron employ advanced analysis techniques like artificial Neural Networks (NN), boosted decision trees (BDT) or matrix element techniques (ME) to combine kinematic characteristics of signal and background events into a single discriminant. These techniques improve the separation of signal to background over the invariant Higgs boson mass distribution which is the most important single variable. Careful validation of all input variables is mandatory for robust results.

Events with neutrinos in the final state are identified using missing transverse energy. The reconstruction of all these variables require excellent performance of all detector components.

2.2 Signal and Background

The Higgs signal is simulated with PYTHIA [6]. The signal cross-sections are normalised to next-to-next-to-leading order (NNLO) calculations [7, 8] and branching ratios from HDECAY [9].

There are many types of background to the Higgs search. An important source of background are multi-jet events (often labeled "QCD background"). This background and the instrumental background due to mis-identified leptons or b jets is either simulated with PYTHIA (only for the CDF $ZH \to \nu\nu b\bar{b}$ analysis) or is taken directly from data, since it is not very well simulated by Monte Carlo. Determining this background from data is done using control samples with no signal content.

Electroweak background processes such as di-boson production, $p\bar{p} \to VV(V = W, Z)$, V+jets or $t\bar{t}$ pair production often dominate at the final stages of the selection; these are simulated using leading order Monte Carlo programs such as PYTHIA, ALPGEN, HERWIG or COMPHEP. The normalisation of these processes is obtained either from data or from NLO calculations.

2.3 Search for $WH \to \ell\nu b\bar{b}$

One of the most sensitive channels for a low Higgs boson mass is the decay $WH \to \ell\nu b\bar{b}$. This final state consists of two b jets from the Higgs boson and a charged lepton ℓ and a neutrino from the W boson. All three leptonic decays of the W boson are analysed at DØ, with the most sensitive being the decays to electrons and muons. Events are selected with one or two b tagged jets an isolated electron or muon and missing transverse energy. The main backgrounds after selection are W+jets and $t\bar{t}$ production. The di-jet invariant mass distribution for events with two b-tags is shown in Figure 2 on the left side. To improve the separation between the signal and the irreducible background a NN is trained which takes a number of kinematic and topological variables as input. The output of this NN is used to extract limits on Higgs production and is shown in Figure 2 on the right side. The analysis uses 1.7 fb $^{-1}$ of recorded data and sets an observed (expected) limit on σ_{95}/σ_{SM}= 9.1(8.5) for a Higgs boson mass $m_H = 115$ GeV/c^2 (where σ_{SM} is the cross section predicted for this process by the Standard Model). A dedicated search for $W^{\pm}H \to \tau^{\pm}\nu b\bar{b}$ with hadronic τ decays has been added at DØ. Using the di-jet mass distribution to separate signal from background an observed (expected) limit on σ_{95}/σ_{SM}= 35.4 (42.1) for a Higgs boson mass $m_H = 115$ GeV/c^2 has been obtained in that channel. At

CDF a similar analysis using 2.7 fb^{-1} of data with a NN discriminant and a combined ME+BDT technique is performed. The analysis sets an observed (expected) limit on σ_{95}/σ_{SM}= 5.0 (5.8) for the NN analysis and σ_{95}/σ_{SM}= 5.8 (5.6) for the ME+BDT analysis for a Higgs boson mass $m_H = 115$ GeV/c^2.

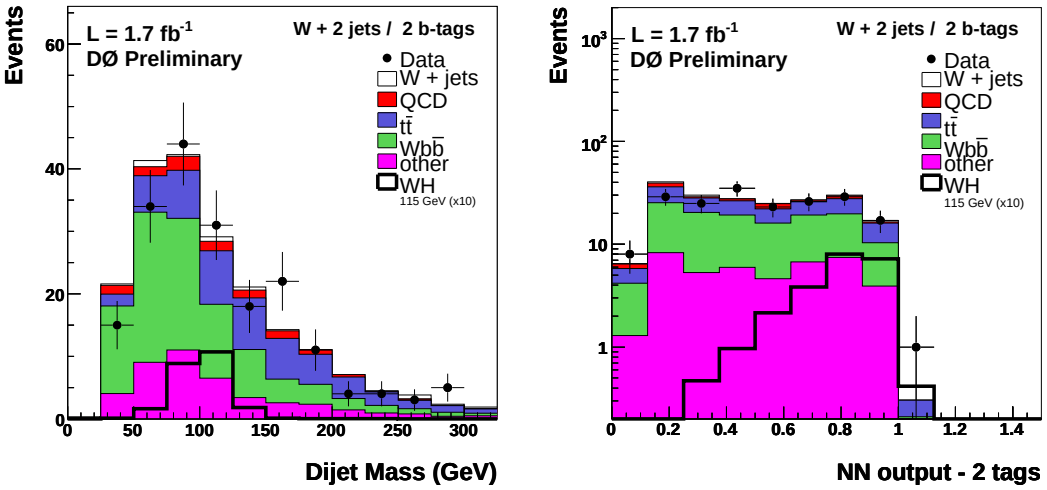

Fig. 2: DØ $WH \to \ell\nu b\bar{b}$ channel: Di-jet invariant mass distribution for events with two b-tags and the NN distribution at the final stage of the selection.

2.4 $ZH \to \nu\nu b\bar{b}$

The channel $ZH \to \nu\nu b\bar{b}$ has very good sensitivity since the branching ratios for $Z \to \nu\nu$ and $H \to b\bar{b}$ decays are large. With the two b-jets being boosted in the transverse direction, the signature for the final state are acoplanar di-jets and large missing transverse energy. Thus is in contrast to most background di-jet events which are expected to be back-to-back in the transverse plane. The main background sources in this search channel are W boson or Z boson production in association with heavy flavour jets, multi-jet events and $t\bar{t}$ pairs.

The basic selection requires at least one (CDF) or two jets (DØ) with a b tag, large missing transverse energy ($E_T^{\rm miss} > 50$ GeV), and a veto on any isolated muon or electron in the event.

In the CDF analysis, the final sample is divided into three samples, one sample with exactly one tight secondary vertex b tag, the second sample with one tight secondary vertex b tag and one tag with the JetProb algorithm and a third sample with two tight secondary vertex b tags. Two NNs are trained one against the dominant QCD background (see Figure 3 on the left side for the second b-tag sample) and one against di-boson and $t\bar{t}$ background (see Figure 3 on the right side for the second sample), which is also used to extract limits on the production cross section.

In the case of DØ, events with two NN b tags are used to construct a BDT for identifying signal events. Asymmetric operating points, one loose and one tight, are chosen for the two b

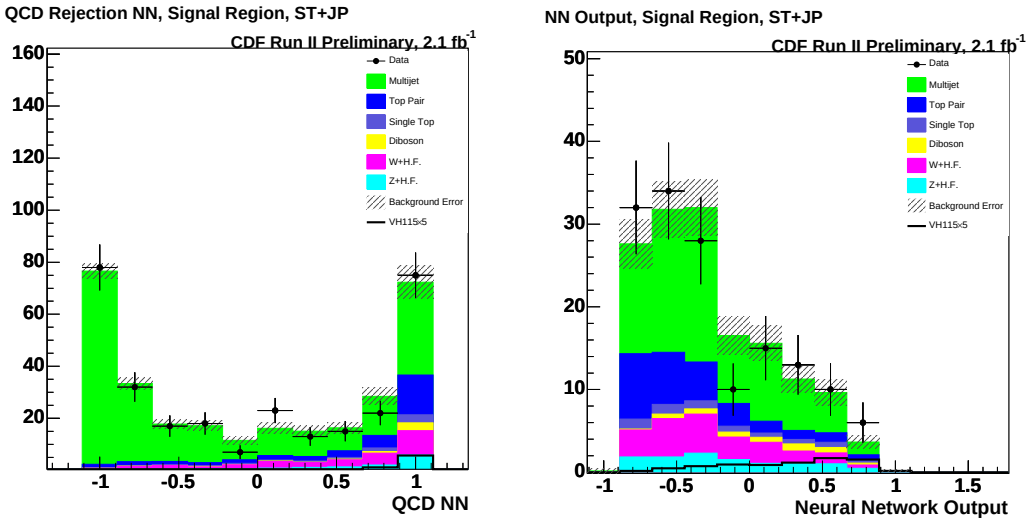

Fig. 3: CDF $ZH \rightarrow \nu\nu b\bar{b}$ channel: NN output distribution to separate against the dominate QCD background (left) and the NN distribution for the remaining backgrounds (right).

tags. The output distributions of the BDT, retrained for every Higgs mass, is shown in Figure 4 on the right side.

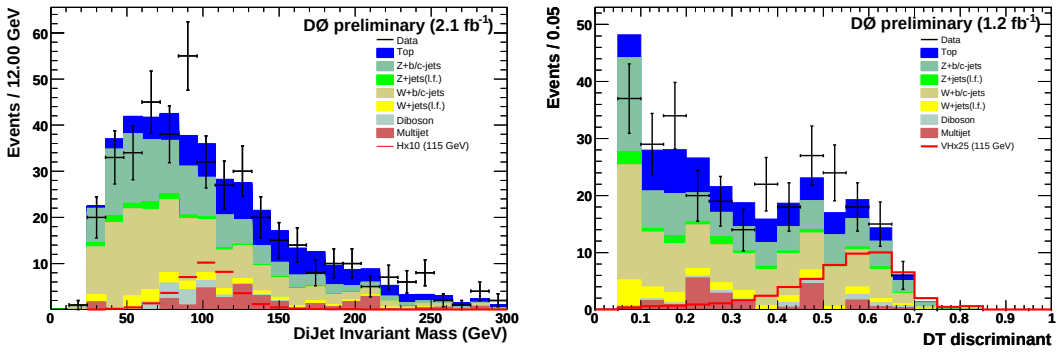

Fig. 4: DØ $ZH \rightarrow \nu\nu b\bar{b}$ channel: Invariant dijet Mass Distribution (left) and output distribution of the BDT variable (right).

To increase the sensitivity of this analysis, WH signal events where the charged lepton has not been identified are also included in the signal definition. This search yields a median observed (expected) upper limit on the $VH(V = W, Z)$ production cross-section of $\sigma_{95}/\sigma_{SM} = 7.9(6.3)$ for CDF and $7.5(8.4)$ for DØ at a Higgs mass of $m_H = 115$ GeV/c^2. The data set for both experiments corresponds to 2.1 fb^{-1} of analyzed data.

2.5 $ZH \to \ell\ell b\bar{b}$

In the $ZH \to \ell\ell b\bar{b}$ channel the Z boson is reconstructed through the decay into two high-p_T isolated muons or electrons. The reconstructed Z and two b-tagged jets are used to select the Higgs signal. The invariant mass of the two leptons is required to be in the Z mass range $70 < m_Z < 110$ GeV/c^2 (DØ) or $76 < m_Z < 106$ GeV/c^2 (CDF). Both experiments require two jets with either one tight b tag or two loose b tags.

The main background sources are Z production in association with heavy jets and $t\bar{t}$ production. ZZ production is an irreducible background, apart from the mass discriminant. CDF trains two separate NNs to reject these two background components. Slices of the output of these NNs, projected on the two axes, is shown in Figure 5. The di-jet mass resolution is improved by training a different NN using E_T^{miss} and the kinematics of both jets. The data set corresponds to an integrated luminosity of 2.4 fb^{-1}. The DØ analysis is performed with 2.3 fb^{-1} of data using a kinematic NN and two NN b tag samples with one tight b tag and two loose b tags.

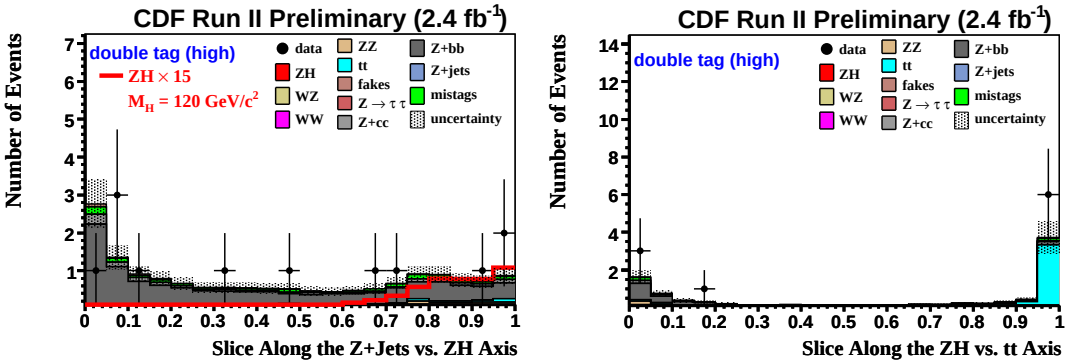

Fig. 5: $ZH \to \ell\ell b\bar{b}$ channel: NN output projection with $y \leq 0.1$ in the Z+Jets vs. ZH projections and $x \geq 0.9$ in the ZH vs. $t\bar{t}$ projection.

These searches yield a median observed (expected) upper limit on the ZH production cross-section of $\sigma_{95}/\sigma_{SM} = 11.6(11.8)$ for CDF and $11.0(12.3)$ for DØ at a Higgs mass of $m_H = 115$ GeV/c^2. Even though the limits are less stringent than for the $ZH \to \nu\nu b\bar{b}$ channel, they still provide an important input to increase the overall sensitivity of the analysis.

2.6 $W \to WW \to \ell\nu\ell\nu$

The dominant decay mode for higher Higgs masses is $H \to WW^{(*)}$. Leptonic decays of the W bosons are therefore used to suppress the QCD background. The signature of the $gg \to H \to WW^{(*)}$ channel is two high-p_T opposite signed isolated leptons with a small azimuthal separation, $\Delta\phi_{\ell\ell}$, due to the spin-correlation between the final-state leptons in the decay of the spin-0 Higgs boson. In contrast, the lepton pairs from background events, mainly WW events, are predominantly back-to-back in $\Delta\phi_{\ell\ell}$. This is shown in Figure 6 (left) for a preselected CDF data sample with zero reconstructed jets.

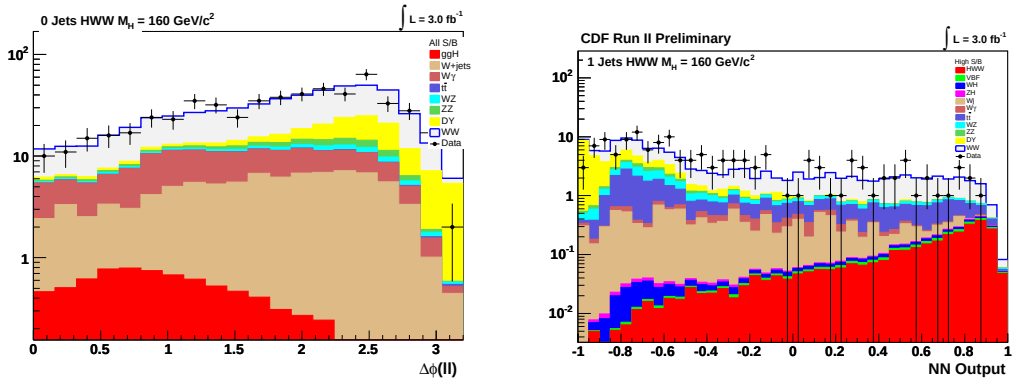

Fig. 6: CDF WW channel: azimuthal angle between the two leptons in the $H \rightarrow WW$ search. Due to spin correlations, the signal is at low $\Delta\phi_{\ell\ell}$, whereas the background is at high $\Delta\phi_{\ell\ell}$.

An additional selection requires $E_T^{\rm miss} > 25$ GeV for CDF and $E_T^{\rm miss} > 20$ GeV for DØ to account for the neutrinos in the final state. DØ defines three final states (e^+e^-, $e^\pm\mu^\mp$, and $\mu^+\mu^-$). CDF separates the $H \rightarrow W^+W^-$ events into five non-overlapping samples, first by separating the events by jet multiplicity (0, 1 or 2), then subdviding the 0 and 1 jet samples in two, one having a low signal/bacgkround (S/B) ratio, the other having a higher one. In these analyses, the final discriminants are neural-network outputs based on several kinematic variables. These include likelihoods constructed from matrix-element probabilities as input to the neural network for CDF and is shown on the right side of Figure 6. The background subtracted NN distribution for DØ is shown in Figure7 on the left side. This distribution has been used to extract median observed (expected) limits on the production cross-section of $\sigma_{95}/\sigma_{SM} = 1.9\,(2.0)$ for $m_H = 165$ GeV/c^2. The obtained limits on the production cross-section as a function of the Higgs boson mass are shown in Figure 7 on the right side. With the NN distributions CDF obtains $\sigma_{95}/\sigma_{SM} = 1.7(1.6)$ for $m_H = 165$ GeV/c^2. The data sets analyzed correspond to an integrated luminosity of 3 fb^{-1} for each experiment.

2.7 $WH \rightarrow WWW^* \rightarrow \ell\nu\ell'\nu q\bar{q}$

In the process $WH \rightarrow WWW^* \rightarrow \ell\nu\ell'\nu q\bar{q}$ the Higgs boson is produced in association with a W boson and subsequently decays into a WW pair. This process is important in the intermediate mass range. The signature is at least two isolated leptons from the W decays with $p_T > 15$ GeV and identical charge. The associated W and one of the two W bosons from the Higgs decay should have the same charge. For the final signal selection DØ used a two-dimensional likelihood based on the invariant mass of the two leptons, the missing transverse energy and their azimuthal angular correlations.

This same-sign charge requirement is very powerful in rejecting background from Z production. The remaining background is either due to di-boson production or due to charge mis-measurements. The rate of charge mis-measurements for muons is determined by comparing the independent charge measurements within the solenoidal and in the toroidal fields of the DØ

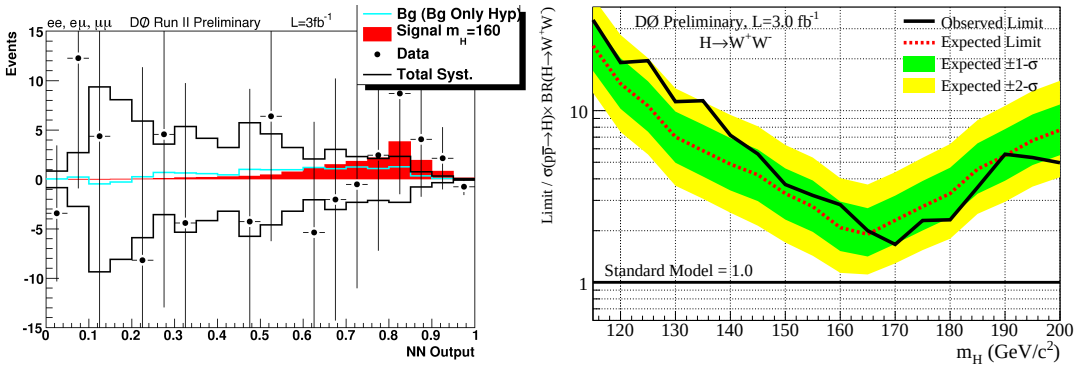

Fig. 7: DØ WW channel: The background subtracted distribution of the NN (left) and the obtained median observed and expected limits on the production cross-section (right).

detector. For electrons the charge mis-measurement rate is determined by comparing the charge measurement from the solenoid with the azimuthal offset between the track and the calorimeter cluster associated to the electron.

The expected cross-section ratio in the mass range 140 GeV/c^2 to 180 GeV/c^2 is $\sigma_{95}/\sigma_{SM} \simeq 20$, i.e. this channel makes a significant contribution at the limit in this mass range.

3 Combined Tevatron Limit

The data of both experiments have been combined using the full set of analyses with luminosities up to 3.0 fb^{-1}. To gain confidence that the final result does not depend on the details of the statistical method applied, several types of combination were performed, using both Modified Frequentist (sometimes called the LEP CL_s method) and Bayesian approaches. The results agree within about 10%. Both methods use Poisson likelihoods and rely on distributions of the final discriminants, e.g. NN output or di-jet mass distributions, not only on event counting.

Systematic uncertainties enter as uncertainties on the expected number of signal and background events, as well as on the shape of the discriminant distributions. The correlations of systematic uncertainties between channels, different background sources, background and signal and between experiments are taken into account. The main sources of systematic uncertainties are, depending on channel, the luminosity and normalisation, the estimates of the multi-jet backgrounds, the input cross-sections used for the MC generated background sources, the higher order corrections (K factors) needed to describe heavy flavour jet production, the jet energy scale, b tagging and lepton identification.

The combinations of results of each single experiment, yield the following ratios of 95% C.L. observed (expected) limits to the SM cross section: 4.2 (3.6) for CDF and 5.3 (4.6) for DØ at $m_H = 115$ GeV/c^2, and 1.8 (1.9) for CDF and 1.7 (2.3) for DØ at $m_H = 170$ GeV/c^2.

The ratios of the 95% C.L. expected and observed limit to the SM cross section are shown in Figure 8 for the combined CDF and DØ analyses on the left side. The observed and median expected values are 1.2 (1.2) at $m_H = 165$ GeV/c^2, 1.0 (1.4) at $m_H = 170$ GeV/c^2 and 1.3

(1.7) at $m_H = 175$ GeV/c^2. On the right side in Figure 8 the 1-CL_S distribution as a function of the Higgs boson mass, which is directly interpreted as the level of exclusion of the search. For instance, both the observed and expected results exclude a Higgs boson with $m_H = 165$ GeV/c^2 at $\approx 92\%$ C.L. The green and yellow bands show the one and two sigma bands for background fluctuations. We exclude at the 95% C.L. the production of a standard model Higgs boson with mass of 170 GeV/c^2.

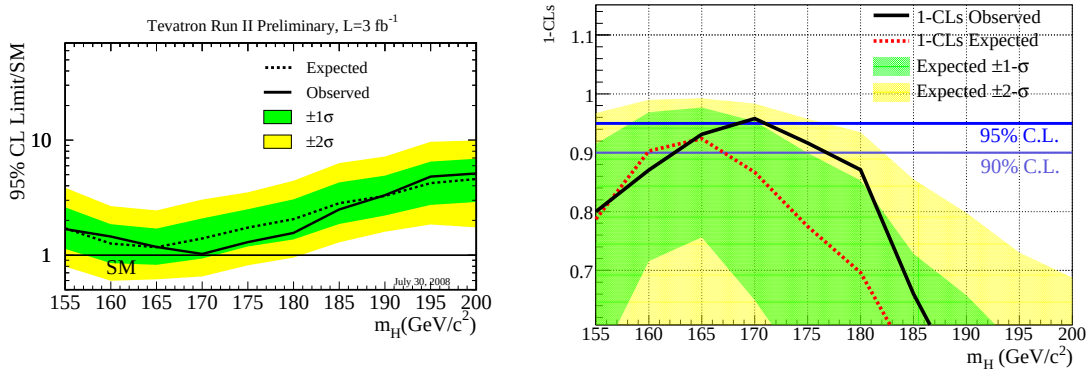

Fig. 8: Expected and observed 95% CL cross-section ratios for the combined CDF and DØ analyses. (status July 2008).

References

[1] LEP Working Group for Higgs boson searches Collaboration, R. Barate *et al.*, Phys. Lett. **B565**, 61 (2003). hep-ex/0306033.

[2] Tevatron Electroweak Working Group Collaboration (2008). http://tevewwg.fnal.gov/.

[3] LEP Electroweak Working Group Collaboration (2008). http://lepewwg.web.cern.ch/LEPEWWG/.

[4] CDF Collaboration, D. E. Acosta *et al.*, Phys. Rev. **D71**, 032001 (2005). hep-ex/0412071.

[5] T. D. M. G. V. M. Abazov, NUCL.INSTRUM.METH.A **552**, 372 (2005).

[6] T. Sjostrand, L. Lonnblad, and S. Mrenna (2001). hep-ph/0108264.

[7] S. Catani, D. de Florian, M. Grazzini, and P. Nason, JHEP **07**, 028 (2003). hep-ph/0306211.

[8] Higgs Working Group Collaboration, K. A. Assamagan *et al.* (2004). hep-ph/0406152.

[9] A. Djouadi, J. Kalinowski, and M. Spira, Comput. Phys. Commun. **108**, 56 (1998). hep-ph/9704448.

Studying the "Underlying Event" at CDF and the LHC

Rick Field[1]
(for the CDF Collaboration)

Department of Physics, University of Florida, Gainesville, Florida, 32611, USA

Abstract

I will report on recent studies of the "underlying event" at CDF using charged particles produced in association with Drell-Yan lepton-pairs in the region of the Z-boson ($70 < M(pair) < 110$ GeV/c^2) in proton-antiproton collisions at 1.96 TeV. The results will be compared with a similar study of the "underlying event" using charged particles produced in association with large transverse momentum jets. The data are corrected to the particle level to remove detector effects and are then compared with several QCD Monte-Carlo models. Some extrapolations of Drell-Yan production to the LHC are also presented.

1. Introduction

In order to find "new" physics at a hadron-hadron collider it is essential to have Monte-Carlo models that simulate accurately the "ordinary" QCD hard-scattering events. To do this one must not only have a good model of the hard scattering part of the process, but also of the beam-beam remnants (BBR) and the multiple parton interactions (MPI). The "underlying event" (*i.e.* BBR plus MPI) is an unavoidable background to most collider observables and a good understanding of it will lead to more precise measurements at the Tevatron and the LHC. Fig. 1.1 illustrates the way the QCD Monte-Carlo models simulate a proton-antiproton collision in which a "hard" 2-to-2 parton scattering with transverse momentum, p_T(hard), has occurred. The resulting event contains particles that originate from the two outgoing partons (*plus initial and final-state radiation*) and particles that come from the breakup of the proton and antiproton (*i.e.* BBR). The "beam-beam remnants" are what is left over after a parton is knocked out of each of the initial two beam hadrons. It is one of the reasons hadron-hadron collisions are more "messy" than electron-positron annihilations and no one really knows how it should be modeled. For the QCD Monte-Carlo models the "beam-beam remnants" are an important component of the "underlying event". Also, multiple parton scatterings contribute to the "underlying event", producing a "hard" component to the "underlying event". Fig. 1.2 shows the way PYTHIA [1] models the "underlying event" in proton-antiproton collision by including multiple parton interactions. In addition to the hard 2-to-2 parton-parton scattering and the "beam-beam remnants", sometimes there are additional "semi-hard" 2-to-2 parton-parton scattering that contribute particles to the "underlying event". The "hard scattering" component consists of the outgoing two jets plus initial and final-state radiation.

As illustrated in Fig. 1.3, the "underlying event" consists of particles that arise from the BBR plus MPI, however, these two components cannot be uniquely separated from particles that come from the initial and final-state radiation. Hence, a study of the "underlying event" inevitably involves a study of the BBR plus MPI plus initial and final-state radiation. As shown in Fig. 1.4, Drell-Yan lepton-pair production provides an excellent place to study the "underlying event". Here one studies the outgoing charged particles (*excluding the lepton pair*) as a function of the lepton-pair invariant mass and as a function of the lepton-pair transverse

[1] This work was done in collaboration with my graduate student Deepak Kar and my former graduate student Craig Group.

momentum. Unlike high p_T jet production for lepton-pair production there is no final-state gluon radiation.

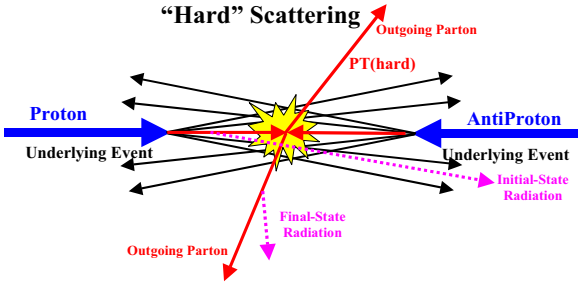

Fig. 1.1. Illustration of the way QCD Monte-Carlo models simulate a proton-antiproton collision in which a "hard" 2-to-2 parton scattering with transverse momentum, P_T(hard), has occurred. The resulting event contains particles that originate from the two outgoing partons (plus initial and final-state radiation) and particles that come from the breakup of the proton and antiproton (*i.e.* "beam-beam remnants"). The "underlying event" is everything except the two outgoing hard scattered "jets" and consists of the "beam-beam remnants" plus initial and final-state radiation. The "hard scattering" component consists of the outgoing two jets plus initial and final-state radiation.

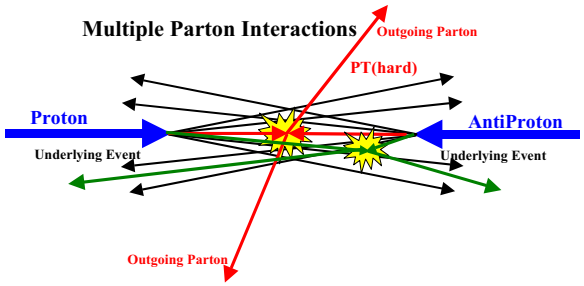

Fig. 1.2. Illustration of the way PYTHIA models the "underlying event" in proton-antiproton collision by including multiple parton interactions. In addition to the hard 2-to-2 parton-parton scattering with transverse momentum, P_T(hard), there is a second "semi-hard" 2-to-2 parton-parton scattering that contributes particles to the "underlying event".

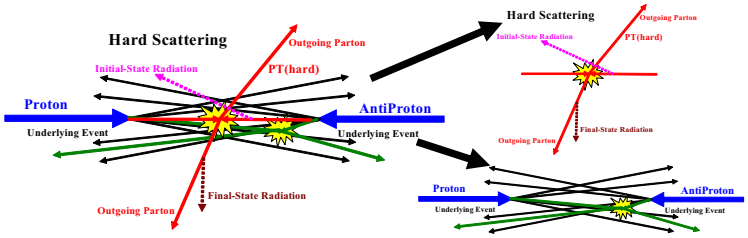

Fig. 1.3. Illustration of the way QCD Monte-Carlo models simulate a proton-antiproton collision in which a "hard" 2-to-2 parton scattering with transverse momentum, P_T(hard), has occurred. The "hard scattering" component of the event consists of particles that result from the hadronization of the two outgoing partons (*i.e.* the initial two "jets") plus the particles that arise from initial and final state radiation (*i.e.* multijets). The "underlying event" consists of particles that arise from the "beam-beam remnants" and from multiple parton interactions.

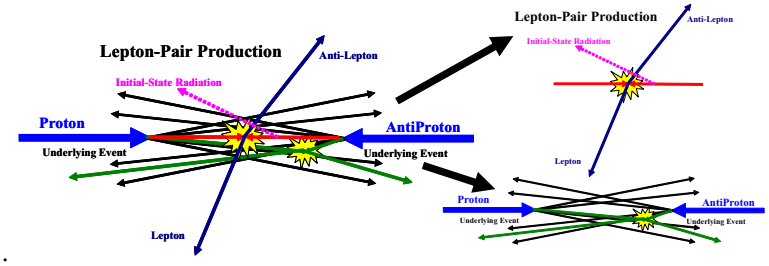

Fig. 1.4. Illustration of the way QCD Monte-Carlo models simulate Drell-Yan lepton-pair production. The "hard scattering" component of the event consists of the two outgoing leptons plus particles that result from initial-state radiation. The "underlying event" consists of particles that arise from the "beam-beam remnants" and from multiple parton interactions.

Hard scattering collider "jet" events have a distinct topology. On the average, the outgoing hadrons "remember" the underlying 2-to-2 hard scattering subprocess. A typical hard scattering event consists of a collection (or burst) of hadrons traveling roughly in the direction of the initial two beam particles and two collections of hadrons (*i.e.* "jets") with large transverse momentum. The two large transverse momentum "jets" are roughly back to back in azimuthal angle. One can use the topological structure of hadron-hadron collisions to study the "underlying event". We use the direction of the leading jet in each event to define four regions of η-ϕ space. As illustrated in Fig. 1.5, the direction of the leading jet, jet#1, in high p_T jet production or the Z-boson in Drell-Yan production is used to define correlations in the azimuthal angle, $\Delta\phi$. The angle $\Delta\phi = \phi - \phi_{jet\#1}$ ($\Delta\phi = \phi - \phi_Z$) is the relative azimuthal angle between a charged particle and the direction of jet#1 (direction of the Z-boson). The "toward" region is defined by $|\Delta\phi| < 60^o$ and $|\eta| < 1$, while the "away" region is $|\Delta\phi| > 120^o$ and $|\eta| < 1$. The two "transverse" regions $60^o < \Delta\phi < 120^o$ and $60^o < -\Delta\phi < 120^o$ are referred to as "transverse 1" and "transverse 2". The overall "transverse" region corresponds to combining the "transverse 1" and "transverse 2" regions. In high p_T jet production, the "toward" and "away" regions receive large contributions from the to the outgoing high p_T jets, while the "transverse" region is perpendicular to the plane of the hard 2-to-2 scattering and is therefore very sensitive to the "underlying event". For Drell-Yan production both the "toward" and the "transverse" region are very sensitive to the "underlying event", while the "away" region receives large contributions from the "away-side" jet from the 2-to-2 processes: $q + \bar{q} \to Z + g$, $q + g \to Z + q$, $\bar{q} + g \to Z + \bar{q}$.

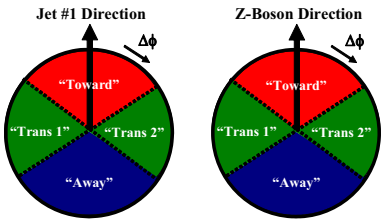

Fig. 1.5. Illustration of correlations in azimuthal angle $\Delta\phi$ relative to (*left*) the direction of the leading jet (highest p_T jet) in the event, jet#1, in high p_T jet production or (*right*) the direction of the Z-boson in Drell-Yan production. The angle $\Delta\phi = \phi - \phi_{jet\#1}$ ($\Delta\phi = \phi - \phi_Z$) is the relative azimuthal angle between charged particles and the direction of jet#1 (Z-boson). The "toward" region is defined by $|\Delta\phi| < 60^o$ and $|\eta| < 1$, while the "away" region is $|\Delta\phi| > 120^o$ and $|\eta| < 1$. The two "transverse" regions $60^o < \Delta\phi < 120^o$ and $60^o < -\Delta\phi < 120^o$ are referred to as "transverse 1" and "transverse 2". Each of the two "transverse" regions have an area in η-ϕ space of $\Delta\eta\Delta\phi = 4\pi/6$. The overall "transverse" region corresponds to combining the "transverse 1" and "transverse 2" regions.

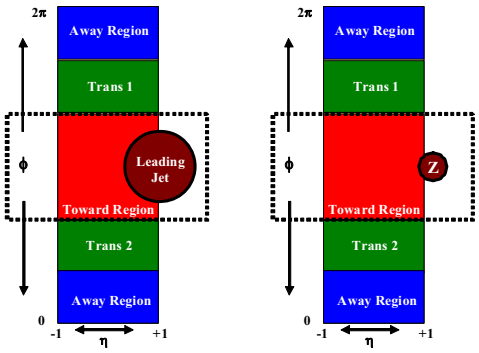

Fig. 1.6. Illustration of correlations in azimuthal angle $\Delta\phi$ relative to (*left*) the direction of the leading jet (highest p_T jet) in the event, jet#1, in high p_T jet production or (*right*) the direction of the Z-boson in Drell-Yan production. The angle $\Delta\phi = \phi - \phi_{jet\#1}$ ($\Delta\phi = \phi - \phi_Z$) is the relative azimuthal angle between charged particles and the direction of jet#1 (Z-boson). The "toward" region is defined by $|\Delta\phi| < 60^o$ and $|\eta| < 1$, while the "away" region is $|\Delta\phi| > 120^o$ and $|\eta| < 1$. The two "transverse" regions $60^o < \Delta\phi < 120^o$ and $60^o < -\Delta\phi < 120^o$ are referred to as "transverse 1" and "transverse 2". We examine charged particles in the range $p_T > 0.5$ GeV/c and $|\eta| < 1$ and $|\eta| < 1$. For high p_T jet production, we require that the leading jet in the event be in the region $|\eta(jet\#1)| < 2$ (referred to as "leading jet" events). For Drell-Yan production we require that invariant mass of the lepton-pair be in the region $81 < M(pair) < 101$ GeV/c^2 with $|\eta(pair)| < 6$ (referred to as "Z-boson" events).

As illustrated in Fig. 1.6, we study charged particles in the range $p_T > 0.5$ GeV/c and $|\eta| < 1$ in the "toward", "away" and "transverse" regions. For high p_T jet production, we require that the leading jet in the event be in the region $|\eta(jet\#1)| < 2$ (referred to as "leading jet" events). The jets are constructed using the MidPoint algorithm ($R = 0.7$, $f_{merge} = 0.75$) . For Drell-Yan production we require that invariant mass of the lepton-pair be in the region $70 < M(pair) < 110$ GeV/c^2 with $|\eta(pair)| < 6$ (referred to as "Z-boson" events).

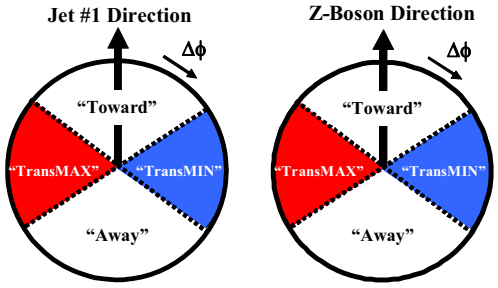

Fig. 1.7. Illustration of correlations in azimuthal angle $\Delta\phi$ relative to the direction of the leading jet (highest p_T jet) in the event, jet#1 for "leading jet" events (*left*) and of correlations in azimuthal angle $\Delta\phi$ relative to the direction of the Z-boson (*right*) in "Z-boson" events. The angle $\Delta\phi$ is the relative azimuthal angle between charged particles and the direction of jet#1 or the Z-boson. On an event by event basis, we define "transMAX" ("transMIN") to be the maximum (minimum) of the two "transverse" regions, $60^o < \Delta\phi < 120^o$ and $60^o < -\Delta\phi < 120^o$. "TransMAX" and "transMIN" each have an area in η-ϕ space of $\Delta\eta\Delta\phi = 4\pi/6$. The overall "transverse" region includes both the "transMAX" and the "transMIN" region.

As shown in Fig. 1.7, for both "leading jet" and "Z-boson" events we define a variety of MAX and MIN "transverse" regions ("transMAX" and "transMIN") which helps separate the "hard component" (initial and final-state radiation) from the "beam-beam remnant" component [2]. MAX (MIN) refer to the "transverse" region containing largest (smallest) number of charged particles or to the region containing the largest (smallest) scalar p_T sum of charged particles. For events with large initial or final-state radiation the "transMAX" region would

contain the third jet in high p_T jet production or the second jet in Drell-Yan production while both the "transMAX" and "transMIN" regions receive contributions from the beam-beam remnants. Thus, the "transMIN" region is very sensitive to the beam-beam remnants, while the "transMAX" minus the "transMIN" (*i.e.* "transDIF") is very sensitive to initial and final-state radiation.

Table 1.1. Observables examined in this analysis as they are defined at the particle level and the detector level. Charged tracks are considered "good" if they pass the track selection criterion. The mean charged particle $<p_T>$ is constructed on an event-by-event basis and then averaged over the events. For the average p_T and the PT_{max} we require that there is at least one charge particle present. The PTsum density is taken to be zero if there are no charged particles present. Particles are considered stable if $c\tau > 10$ mm (*i.e.* K_s, Λ, Σ, Ξ, and Ω are kept stable) .

Observable	Particle Level	Detector level
$dN/d\eta d\phi$	Number of stable charged particles per unit η-ϕ ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$)	Number of "good" tracks per unit η-ϕ ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$)
$dPT/d\eta d\phi$	Scalar p_T sum of stable charged particles per unit η-ϕ ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$)	Scalar p_T sum of "good" tracks per unit η-ϕ ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$)
$<p_T>$	Average p_T of stable charged particles ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$) Require at least 1 charged particle	Average p_T of "good" tracks ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$) Require at least 1 "good" track
PT_{max}	Maximum p_T stable charged particle ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$) Require at least 1 charged particle	Maximum p_T "good" charged tracks ($p_T > 0.5$ GeV/c, $\|\eta\| < 1$) Require at least 1 "good" track
"Jet"	MidPoint algorithm R = 0.7 f_{merge} = 0.75 applied to stable particles	MidPoint algorithm R = 0.7 f_{merge} = 0.75 applied to calorimeter cells

The CDF data are corrected to the particle level to remove detector effects. Table 1.1 shows the observables that are considered in this analysis as they are defined at the particle level and detector level. Since we will be studying regions in η-ϕ space with different areas, we will construct densities by dividing by the area. For example, the number density, $dN/d\eta d\phi$, corresponds the number of charged particles per unit η-ϕ and the PT_{sum} density, $dPT/d\eta d\phi$, corresponds the amount of charged scalar p_T sum per unit η-ϕ. The corrected observables are then compared with QCD Monte-Carlo predictions at the particle level (*i.e.* generator level).

2. QCD Monte-Carlo Model Tunes

PYTHIA Tune A was determined by fitting the CDF Run 1 "underlying event" data [3] and, at that time, we did not consider the "Z-boson" data. Tune A does not fit the CDF Run 1 Z-boson p_T distribution very well [4]. PYTHIA Tune AW fits the Z-boson p_T distribution as well as the "underlying event" at the Tevatron [5]. For "leading jet" production Tune A and Tune AW are nearly identical. Table 2.1 shows the parameters for several PYTHIA 6.2 tunes. PYTHIA Tune DW is very similar to Tune AW except PARP(67) = 2.5, which is the preferred value determined by DØ in fitting their dijet $\Delta\phi$ distribution [6]. PARP(67) sets the high p_T scale for initial-state radiation in PYTHIA. It determines the maximal parton virtuality allowed in time-like showers. Tune DW and Tune DWT are identical at 1.96 TeV, but Tune DW and DWT extrapolate

differently to the LHC. Tune DWT uses the ATLAS energy dependence, PARP(90) = 0.16, while Tune DW uses the Tune A value of PARP(90) = 0.25. All these tunes use CTEQ5L.

The first 9 parameters in Table 2.1 tune the multiple parton interactions (MPI). PARP(62), PARP(64), and PARP(67) tune the initial-state radiation and the last three parameters set the intrinsic k_T of the partons within the incoming proton and antiproton.

Table 2.1. Parameters for several PYTHIA 6.2 tunes. Tune A is the CDF Run 1 "underlying event" tune. Tune AW and DW are CDF Run 2 tunes which fit the existing Run 2 "underlying event" data and fit the Run 1 Z-boson p_T distribution. The ATLAS Tune is the tune used in the ATLAS TRD. Tune DWT use the ATLAS energy dependence for the MPI, PARP(90). The first 9 parameters tune the multiple parton interactions. PARP(62), PARP(64), and PARP(67) tune the initial-state radiation and the last three parameters set the intrinsic k_T of the partons within the incoming proton and antiproton.

Parameter	Tune A	Tune AW	Tune DW	Tune DWT	ATLAS
PDF	CTEQ5L	CTEQ5L	CTEQ5L	CTEQ5L	CTEQ5L
MSTP(81)	1	1	1	1	1
MSTP(82)	4	4	4	4	4
PARP(82)	2.0	2.0	1.9	1.9409	1.8
PARP(83)	0.5	0.5	0.5	0.5	0.5
PARP(84)	0.4	0.4	0.4	0.4	0.5
PARP(85)	0.9	0.9	1.0	1.0	0.33
PARP(86)	0.95	0.95	1.0	1.0	0.66
PARP(89)	1800	1800	1800	1960	1000
PARP(90)	0.25	0.25	0.25	0.16	0.16
PARP(62)	1.0	1.25	1.25	1.25	1.0
PARP(64)	1.0	0.2	0.2	0.2	1.0
PARP(67)	4.0	4.0	2.5	2.5	1.0
MSTP(91)	1	1	1	1	1
PARP(91)	1.0	2.1	2.1	2.1	1.0
PARP(93)	5.0	15.0	15.0	15.0	5.0

Table 2.2. Shows the computed value of the multiple parton scattering cross section for the various PYTHIA 6.2 tunes.

Tune	σ(MPI) at 1.96 TeV	σ(MPI) at 14 TeV
A, AW	309.7 mb	484.0 mb
DW	351.7 mb	549.2 mb
DWT	351.7 mb	829.1 mb
ATLAS	324.5 mb	768.0 mb

Table 2.2 shows the computed value of the multiple parton scattering cross section for the various tunes. The multiple parton scattering cross section (divided by the total inelastic cross section) determines the average number of multiple parton collisions per event.

JIMMY [7] is a multiple parton interaction model which can be added to HERWIG [8] to improve agreement with the "underlying event" observables. To compare with the "Z-boson" data we have constructed a HERWIG (with JIMMY MPI) tune with JMUEO = 1, PTJIM = 3.6 GeV/c, JMRAD(73) = 1.8, and JMRAD(91) = 1.8.

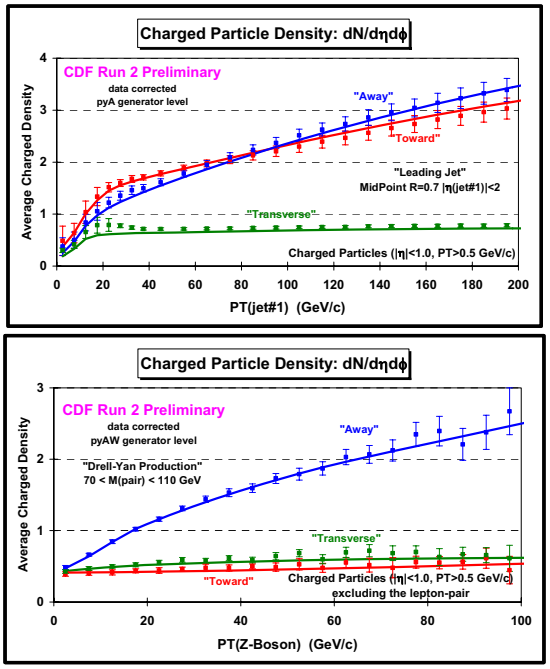

Fig. 3.1. CDF data at 1.96 TeV on the density of charged particles, dN/dηdϕ, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "leading jet" (*top*) and "Z-boson" (*bottom*) events as a function of the leading jet p_T and $p_T(Z)$, respectively, for the "toward", "away", and "transverse" regions. The data are corrected to the particle level and are compared with PYTHIA Tune A and Tune AW, respectively, at the particle level (*i.e.* generator level).

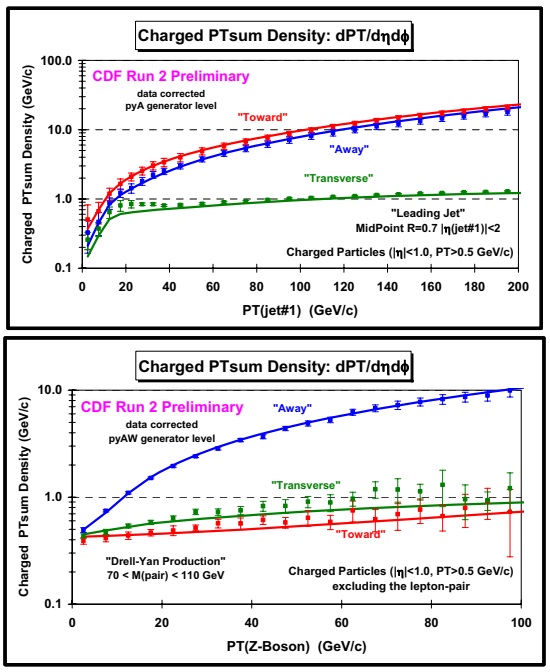

Fig. 3.2. CDF data at 1.96 TeV on the *scalar* PTsum density of charged particles, dPT/dηdϕ, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ and "leading jet" (*top*) and "Z-Boson" (*bottom*) events as a function of the leading jet p_T and $p_T(Z)$, respectively, for the "toward",

"away", and "transverse" regions. The data are corrected to the particle level and are compared with PYTHIA Tune A and Tune AW, respectively, at the particle level (*i.e.* generator level).

3. CDF results

3.1 "Leading Jet" and "Z-Boson" Topologies

Fig. 3.1 and Fig. 3.2 show the data on the density of charged particles and the *scalar* PTsum density, respectively, for the "toward", "away", and "transverse" regions for "leading jet" and "Z-boson" events. For "leading jet" events the densities are plotted as a function of the leading jet p_T and for "Z-boson" events there are plotted versus $p_T(Z)$. The data are corrected to the particle level and are compared with PYTHIA Tune A ("leading jet") and Tune AW ("Z-boson") at the particle level (*i.e.* generator level). For "leading jet" events at high p_T(jet#1) the densities in the "toward" and "away" regions are much larger than in the "transverse" region because of the "toward-side" and "away-side" jets. At small p_T(jet#1) the "toward", "away", and "transverse" densities become equal and go to zero as p_T(jet#1) goes to zero. As the leading jet transverse momentum becomes small all three regions are populated by the underlying event and if the leading jet has no transverse momentum then there are no charged particles anywhere. There are a lot of low transverse momentum jets and for p_T(jet#1) < 30 GeV/c and the leading jet is not always the jet resulting from the hard 2-to-2 scattering. This produces a "bump" in the "transverse" density in the range where the "toward", "away", and "transverse" densities become similar in size. For "Z-boson" events the "toward" and "transverse" densities are both small and almost equal. The "away" density is large due to the "away-side" jet. The "toward", "away", and "transverse" densities become equal as $p_T(Z)$ goes to zero, but unlike the "leading jet" case the densities do not vanish at $p_T(Z) = 0$. For "Z-boson" events with $p_T(Z) = 0$ the hard scale is set by the Z-boson mass, whereas in "leading jet" events the hard scale goes to zero as the transverse momentum of the leading jet goes to zero.

Fig. 3.3 compares the data for "leading jet" events with the data for "Z-boson" events for the density of charged particles in the "transverse" region. The data are compared with PYTHIA Tune A ("leading jet") , Tune AW ("Z-boson"), and HERWIG (without MPI). For large p_T(jet#1) the "transverse" densities are similar for "leading jet" and "Z-boson" events as one would expect. HERWIG (without MPI) does not produce enough activity in the "transverse" region for either process. HERWIG (without MPI) disagrees more with the "transverse" region of "Z-boson" events than it does with the "leading jet" events. This is because there is no final-state radiation in "Z-boson" production so that the lack of MPI becomes more evident.

Fig. 3.4 compares the data for "leading jet" events with the data for "Z-boson" events for the average charged particle p_T in the "transverse" region. The data are compared with PYTHIA Tune A ("leading jet") , Tune AW ("Z-boson"), and HERWIG (without MPI). MPI provides a "hard" component to the "underlying event" and for HERWIG (without MPI) the p_T distributions in the "transverse" region for both processes are too "soft", resulting in an average p_T that is too small.

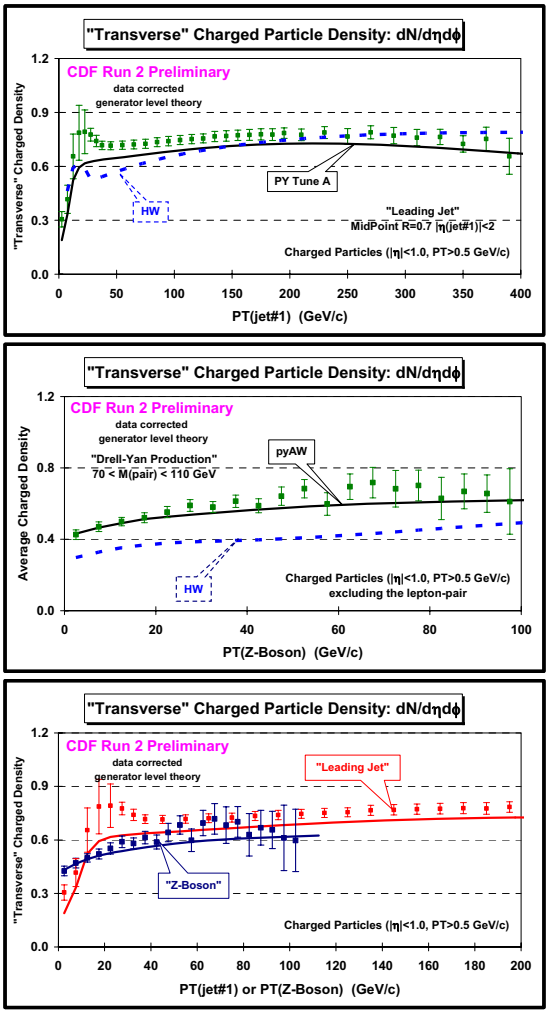

Fig. 3.3. (*top*) Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "leading jet" events as a function of the leading jet p_T in the "transverse" region compared with HERWIG (without MPI) and PYTHIA Tune A at the particle level (*i.e.* generator level). (*middle*) Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of the leading jet $p_T(Z)$ in the "transverse" region compared with HERWIG (without MPI) and PYTHIA Tune AW at the particle level (*i.e.* generator level). (*bottom*) Data on the density of charged particles for "leading jet" and "Z-boson" events as a function of the leading jet p_T and $p_T(Z)$, respectively, for the "transverse" region compared with PYTHIA Tune A ("leading jet") and Tune AW ("Z-boson").

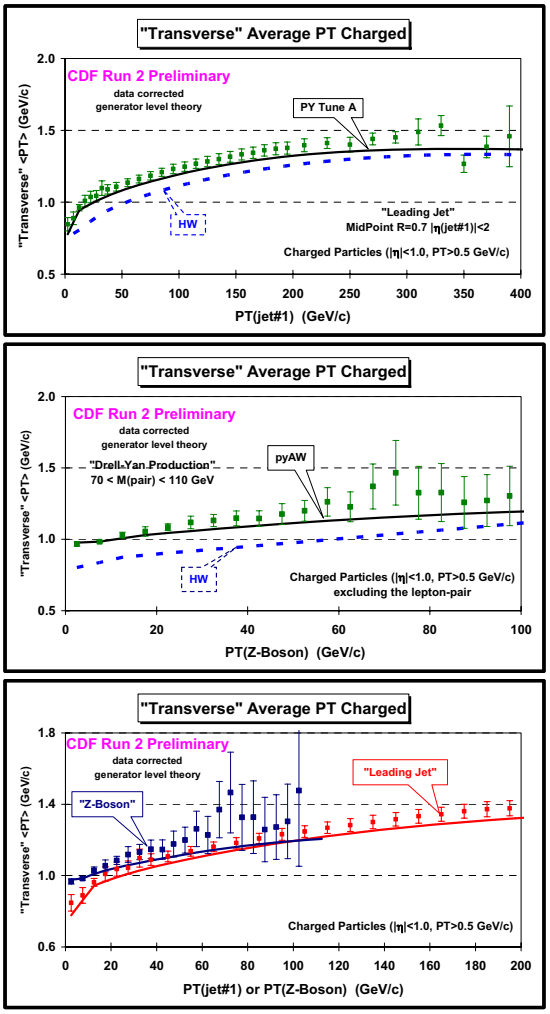

Fig. 3.4. (*top*) Data corrected to the particle level at 1.96 TeV on the average charged particle transverse momentum, $\langle p_T \rangle$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "leading jet" events as a function of the leading jet p_T in the "transverse" region compared with HERWIG (without MPI) and PYTHIA Tune A at the particle level (*i.e.* generator level). (*middle*) Data corrected to the particle level at 1.96 TeV on the average charged particle transverse momentum, $\langle p_T \rangle$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of the leading jet $p_T(Z)$ in the "transverse" region compared with HERWIG (without MPI) and PYTHIA Tune AW at the particle level (*i.e.* generator level). (*bottom*) Data on the average charged particle transverse momentum for "leading jet" and "Z-boson" events as a function of the leading jet p_T and $p_T(Z)$, respectively, for the "transverse" region compared with PYTHIA Tune A ("leading jet") and Tune AW ("Z-boson").

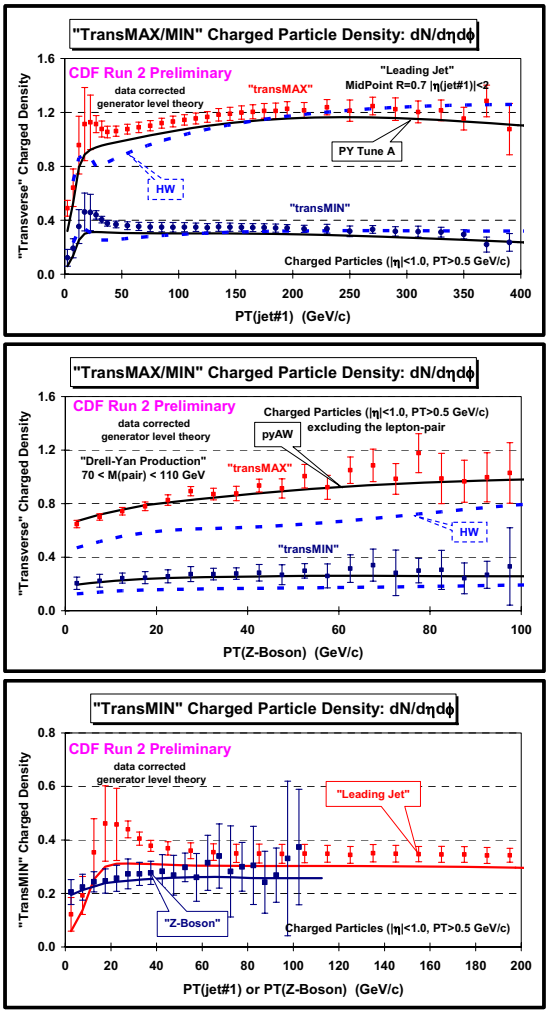

Fig. 3.5. (*top*) Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "leading jet" events as a function of the leading jet p_T for the "transMAX" and "transMIN" regions compared with HERWIG (without MPI) and PYTHIA Tune A at the particle level (*i.e.* generator level). (*middle*) Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of the leading jet $p_T(Z)$ for the "transMAX" and "transMIN" regions compared with HERWIG (without MPI) and PYTHIA Tune AW at the particle level (*i.e.* generator level). (*bottom*) Data on the density of charged particles for "leading jet" and "Z-boson" events as a function of the leading jet p_T and $p_T(Z)$, respectively, for the "transMIN" region compared with PYTHIA Tune A ("leading jet") and Tune AW ("Z-boson").

Fig. 3.5 compares the data for "leading jet" events with the data for "Z-boson" events for the density of charged particles for the "transMAX" and "transMIN" regions. The data are compared with PYTHIA Tune A ("leading jet") , Tune AW ("Z-boson"), and HERWIG (without MPI).

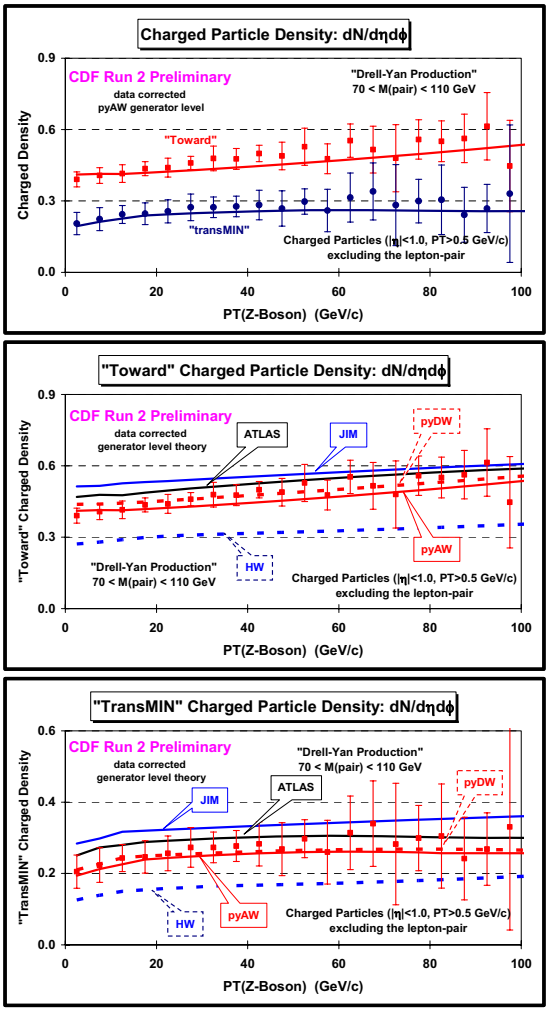

Fig. 3.6. Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of $p_T(Z)$, in the "toward" and "transMIN" regions. (*top*) Data in the "toward" and "transMIN" regions are compared with PYTHIA Tune AW. (*middle*) Data in the "toward" region are compared with HERWIG (without MPI), HERWIG (with JIMMY MPI), and three PYTHIA MPI tunes (AW, DW, ATLAS). (*middle*) Data for the "transMIN" region are compared with HERWIG (without MPI), HERWIG (with JIMMY MPI), and three PYTHIA MPI tunes (AW, DW, ATLAS).

3.2 The "Underlying Event" in Drell-Yan Production

The most sensitive regions to the "underlying event" in Drell-Yan production are the "toward" and the "transMIN" regions, since these regions are less likely to receive contributions from initial-state radiation. Fig. 3.6 and Fig. 3.7 show the data for "Z-boson" events for the density of charged particles and the *scalar* PTsum density, respectively, in the "toward" and "transMIN" regions. The data are compared with PYTHIA Tune AW, Tune DW, the PYTHIA ATLAS tune. HERWIG (without MPI), and HERWIG (with JIMMY MPI). The densities are smaller in the "transMIN" region than in the "toward" region and this is described well by

PYTHIA Tune AW. Comparing HERWIG (without MPI) with HERWIG (with JIMMY MPI) clearly shows the importance of MPI in these regions. Tune AW and Tune DW are very similar. The ATLAS tune and HERWIG (with JIMMY MPI) agree with Tune AW for the *scalar* PTsum density in the "toward" and "transMIN" regions. However, both the ATLAS tune and HERWIG (with JIMMY MPI) produce too much charged particle density in these regions. The ATLAS tune and HERWIG (with JIMMY MPI) fit the PTsum density, but they do so by producing too many charged particles (i.e. they both have to "soft" of a p_T spectrum in these regions). This can be seen clearly in Fig. 3.8 which shows the data for "Z-boson" events on the average charged particle p_T and the average maximum charged particle p_T, in the "toward" region compared with the QCD Monte-Carlo models.

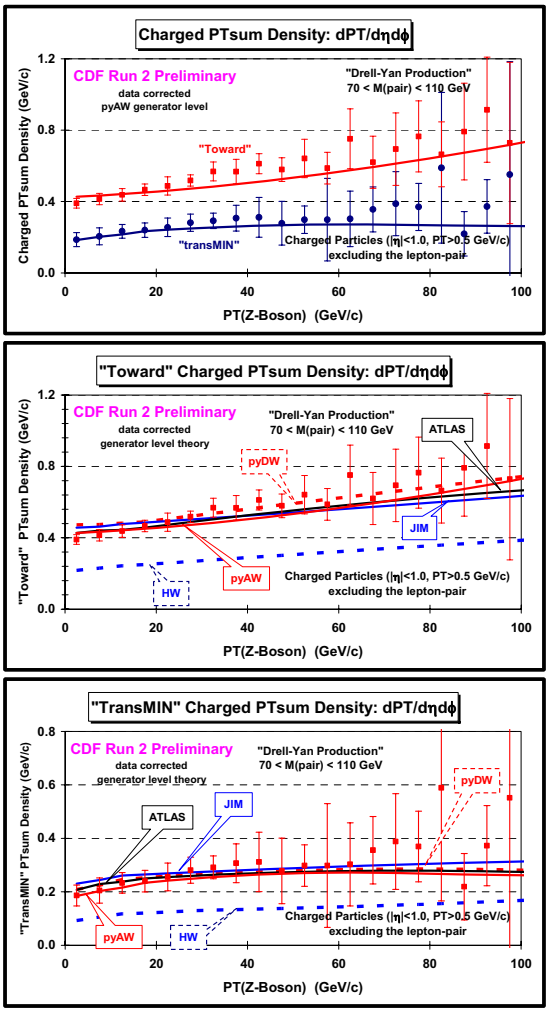

Fig. 3.7. Data corrected to the particle level at 1.96 TeV on the *scalar* charged particle PTsum density, dPT/dηdϕ, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of $p_T(Z)$, in the "toward" and "transMIN" regions. (*top*) Data for the "toward" and "transMIN" regions are compared with PYTHIA Tune AW. (*middle*) Data for the "toward" region are compared with HERWIG (without MPI), HERWIG (with JIMMY MPI), and three PYTHIA MPI tunes (AW, DW, ATLAS). (*middle*) Data for the "transMIN" region are compared with HERWIG (without MPI), HERWIG (with JIMMY MPI), and three PYTHIA MPI tunes (AW, DW, ATLAS).

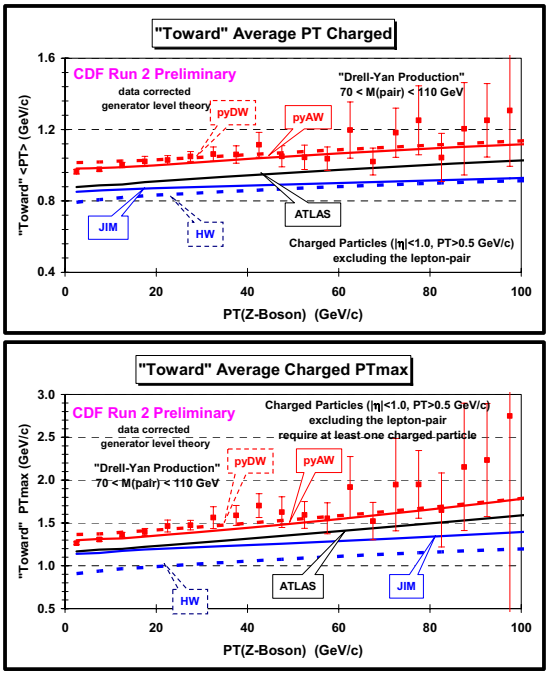

Fig. 3.8. Data corrected to the particle level at 1.96 TeV on the charged particle average transverse momentum, $\langle p_T \rangle$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ (*top*) and average maximum charged particle transverse momentum, <PTmax>, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ (require at least one charged particle) (*bottom*) for "Z-boson" events as a function of $p_T(Z)$, in the "toward" region compared with HERWIG (without MPI), HERWIG (with JIMMY MPI), and three PYTHIA MPI tunes (AW, DW, ATLAS).

3.3 Extrapolating Drell-Yan Production to the LHC

Fig. 3.9 shows the extrapolation of PYTHIA Tune DWT and HERWIG (without MPI) for the density of charged particles and the average transverse momentum of charged particles in the "towards" region of "Z-boson" production to 10 TeV (LHC10) and to 14 TeV (LHC14). For HERWIG (without MPI) the "toward" region of "Z-boson" production does not change much in going from the Tevatron to the LHC. Models with multiple-parton interactions like PYTHIA Tune DWT predict that the "underlying event" will become much more active (with larger $\langle p_T \rangle$) at the LHC.

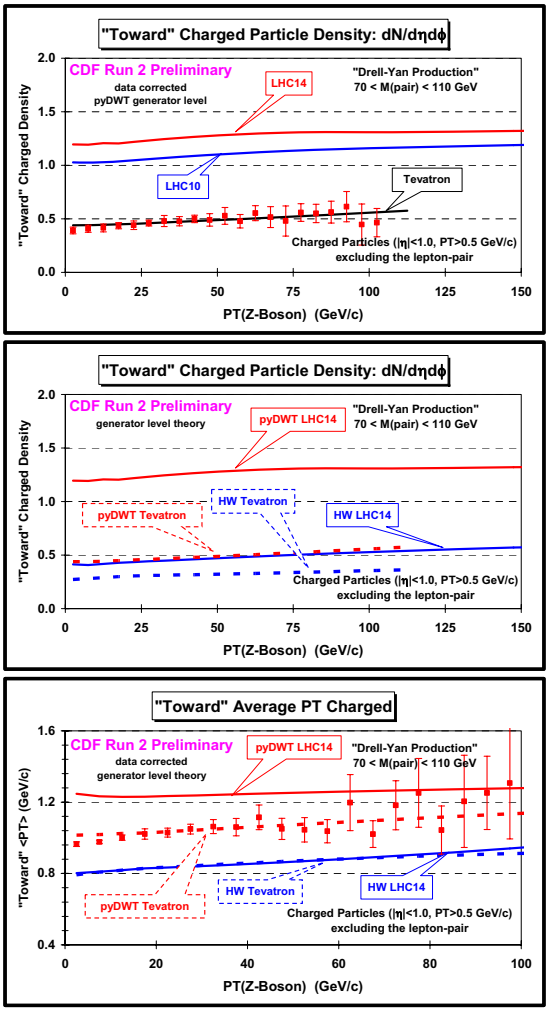

Fig. 3.9. (*top*) Data corrected to the particle level at 1.96 TeV on the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of $p_T(Z)$, in the "toward" region compared with PYTHIA Tune DWT at 1.96 TeV (Tevatron), 10 TeV (LHC10), and 14 TeV (LHC14). (*middle*) Predictions of HERWIG (without MPI) and PYTHIA Tune DWT for the density of charged particles, $dN/d\eta d\phi$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of $p_T(Z)$, in the "toward" region at 1.96 TeV (Tevatron) and 14 TeV (LHC14). (*bottom*) Data corrected to the particle level at 1.96 TeV on the average charged particle transverse momentum. $\langle p_T \rangle$, with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events as a function of $p_T(Z)$, for the "toward" region compared with HERWIG (without MPI) and PYTHIA Tune DWT at 1.96 TeV (Tevatron) and 14 TeV (LHC14).

3.4 $\langle p_T \rangle$ versus the Multiplicity: "Min-Bias" and "Z-boson" Events

The total proton-antiproton cross section is the sum of the elastic and inelastic components, $\sigma_{tot} = \sigma_{EL} + \sigma_{IN}$. The inelastic cross section consists of three terms; single diffraction, double-diffraction, and everything else (referred to as the "hard core"), $\sigma_{IN} = \sigma_{SD} + \sigma_{DD} + \sigma_{HC}$. For elastic scattering neither of the beam particles breaks apart (*i.e.* color singlet exchange). For single and double diffraction one or both of the beam particles are excited into a high mass color

singlet state (*i.e.* N^* states) which then decays. Single and double diffraction also corresponds to color singlet exchange between the beam hadrons. When color is exchanged the outgoing remnants are no longer color singlets and one has a separation of color resulting in a multitude of quark-antiquark pairs being pulled out of the vacuum. The "hard core" component, σ_{HC}, involves color exchange and the separation of color. However, the "hard core" contribution has both a "soft" and "hard" component. Most of the time the color exchange between partons in the beam hadrons occurs through a soft interaction (*i.e.* no high transverse momentum) and the two beam hadrons "ooze" through each other producing lots of soft particles with a uniform distribution in rapidity and many particles flying down the beam pipe. Occasionally there is a hard scattering among the constituent partons producing outgoing particles and "jets" with high transverse momentum.

Minimum bias (*i.e.* "min-bias") is a generic term which refers to events that are selected with a "loose" trigger that accepts a large fraction of the inelastic cross section. All triggers produce some bias and the term "min-bias" is meaningless until one specifies the precise trigger used to collect the data. The CDF "min-bias" trigger consists of requiring at least one charged particle in the forward region $3.2 < \eta < 5.9$ and simultaneously at least one charged particle in the backward region $-5.9 < \eta < -3.2$. Monte-Carlo studies show that the CDF "min-bias" collects most of the σ_{HC} contribution plus small amounts of single and double diffraction.

Minimum bias collisions are a mixture of hard processes (perturbative QCD) and soft processes (non-perturbative QCD) and are, hence, very difficult to simulate. Min-bias collisions contain soft "beam-beam remnants", hard QCD 2-to-2 parton-parton scattering, and multiple parton interactions (soft & hard). To correctly simulate min-bias collisions one must have the correct mixture of hard and soft processes together with a good model of the multiple-parton interactions. The first model that came close to correctly modeling min-bias collisions at CDF was PYTHIA Tune A. Tune A was not tuned to fit min-bias collisions. It was tuned to fit the activity in the "underlying event" in high transverse momentum jet production [3]. However, PYTHIA uses the same p_T cut-off for the primary hard 2-to-2 parton-parton scattering and for additional multiple parton interactions. Hence, fixing the amount of multiple parton interactions (*i.e.* setting the p_T cut-off) allows one to run the hard 2-to-2 parton-parton scattering all the way down to $p_T(\text{hard}) = 0$ without hitting a divergence. For PYTHIA the amount of hard scattering in min-bias is, therefore, related to the activity of the "underlying event" in hard scattering processes. Neither HERWIG (without MPI) or HERWIG (with JIMMY MPI) can be used to describe "min-bias" events since they diverge as $p_T(\text{hard})$ goes to zero.

Fig. 3.10 shows the new CDF "min-bias" data presented at this conference by Niccolo' Moggi [9]. The data are corrected to the particle level at 1.96 TeV and show the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.4$ GeV/c and $|\eta| < 1$. The data are compared with PYTHIA Tune A, the PYTHIA ATLAS tune, and PYTHIA Tune A without MPI (pyAnoMPI). This is an important observable. The rate of change of $\langle p_T \rangle$ versus charged multiplicity is a measure of the amount of hard versus soft processes contributing to min-bias collisions and it is sensitive to the modeling of the multiple-parton interactions [10]. If only the soft "beam-beam" remnants contributed to min-bias collisions then $\langle p_T \rangle$ would not depend on charged multiplicity. If one has two processes contributing, one soft ("beam-beam remnants") and one hard (hard 2-to-2 parton-parton scattering), then demanding large multiplicity will preferentially select the hard process and lead to a high $\langle p_T \rangle$. However, we see that with only these two processes $\langle p_T \rangle$ increases much too rapidly as a function of multiplicity (see pyAnoMPI). Multiple-parton interactions provides another mechanism for producing large

multiplicities that are harder than the "beam-beam remnants", but not as hard as the primary 2-to-2 hard scattering. PYTHIA Tune A gives a fairly good description of the $\langle p_T \rangle$ versus multiplicity, although not perfect. PYTHIA Tune A does a better job describing the data than the ATLAS tune. Both Tune A and the ATLAS tune include multiple-parton interactions, but with different choices for the color connections [11].

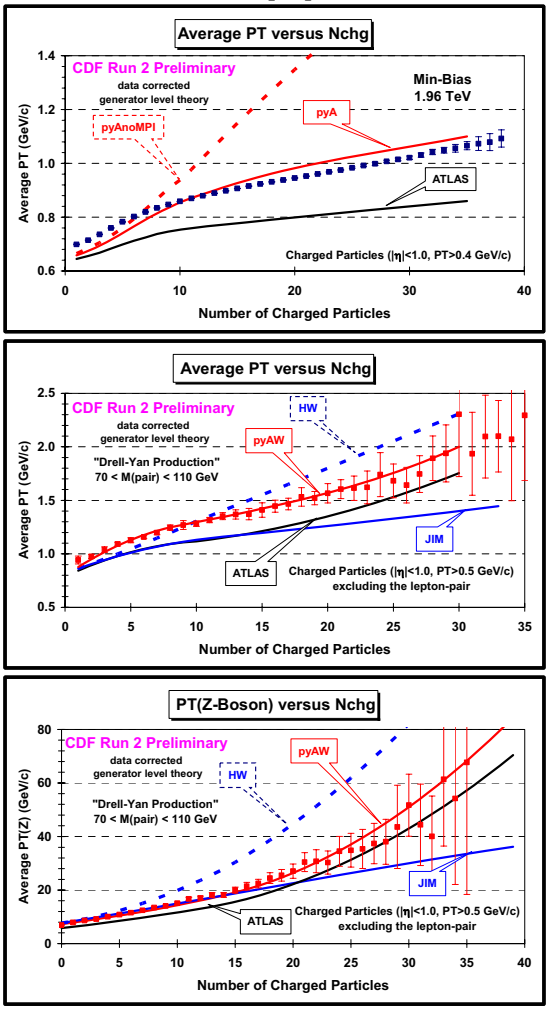

Fig. 3.10. (*top*) CDF "Min-Bias" data corrected to the particle level at 1.96 TeV on the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.4$ GeV/c and $|\eta| < 1$ from Ref. 14. The data are compared with PYTHIA Tune A, the PYTHIA ATLAS tune, and PYTHIA Tune A without MPI (pyAnoMPI). (*middle*) Data corrected to the particle level at 1.96 TeV on the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events. (*bottom*) Data corrected to the particle level at 1.96 TeV on the average p_T of the Z-boson versus the multiplicity for charged particles with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events. The "Z-boson" data are compared with PYTHIA Tune AW, the PYTHIA ATLAS tune, HERWIG (without MPI), and HERWIG (with JIMMY MPI).

Fig. 3.9 also shows the data at 1.96 TeV on the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events from this analysis. HERWIG (without MPI) predicts the $\langle p_T \rangle$ to rise too rapidly as the multiplicity increases. This is similar to the pyAnoMPI behavior in "min-bias" collisions. For HERWIG

(without MPI) large multiplicities come from events with a high p_T Z-boson and hence a large p_T "away-side" jet. This can be seen clearly in Fig. 3.10 which also shows the average p_T of the Z-boson versus the charged multiplicity. Without MPI the only way of getting large multiplicity is with high $p_T(Z)$ events. For the models with MPI one can get large multiplicity either from high $p_T(Z)$ events or from MPI and hence $<P_T(Z)>$ does not rise as sharply with multiplicity in accord with the data. PYTHIA Tune AW describes the data "Z-boson" fairly well.

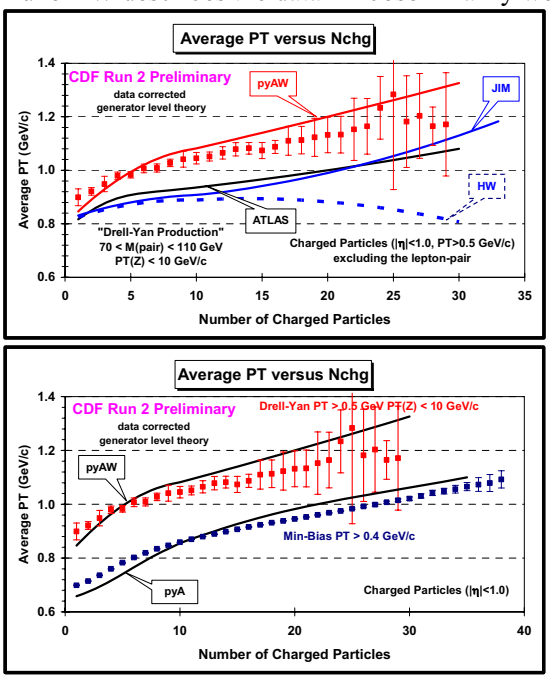

Fig. 3.11. (*top*) Data corrected to the particle level at 1.96 TeV on the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events in which $p_T(Z) < 10$ GeV/c. The data are compared with PYTHIA Tune AW, the PYTHIA ATLAS tune, HERWIG (without MPI), and HERWIG (with JIMMY MPI). (*bottom)* Comparison of the average p_T of charged particles versus the charged multiplicity for "Min-Bias" events from Ref. 14 with the "Z-boson" events with $p_T(Z) < 10$ GeV/c from this analysis. The "Min-Bias" data require $p_T > 0.4$ GeV/c and are compared with PYTHIA Tune A, while the "Z-boson" data require $p_T > 0.5$ GeV/c and are compared with PYTHIA Tune AW.

Fig. 3.11 shows the data at 1.96 TeV on the average p_T of charged particles versus the multiplicity for charged particles with $p_T > 0.5$ GeV/c and $|\eta| < 1$ for "Z-boson" events in which $p_T(Z) < 10$ GeV/c. We see that $<p_T>$ still increases as the multiplicity increases although not as fast. If we require $p_T(Z) < 10$ GeV/c, then HERWIG (without MPI) predicts that the $<p_T>$ decreases slightly as the multiplicity increases. This is because without MPI and without the high p_T "away-side" jet which is suppressed by requiring low $p_T(Z)$, large multiplicities come from events with a lot of initial-state radiation and the particles coming from initial-state radiation are "soft". PYTHIA Tune AW describes the behavior of $<p_T>$ versus the multiplicity fairly well even when we select $p_T(Z) < 10$ GeV/c.

Fig. 3.11 also shows a comparison of the average p_T of charged particles versus the charged multiplicity for "min-bias" events [9] with the "Z-boson" events with $p_T(Z) < 10$ GeV/c. There is no reason for the "min-bias" data to agree with the "Z-boson" events with $p_T(Z) < 10$ GeV/c. However, they are remarkably similar and described fairly well by PYTHIA Tune A and Tune

AW, respectively. This strongly suggests that MPI are playing an important role in both these processes.

4. Summary & Conclusions

Observables that are sensitive to the "underlying event" in high transverse momentum jet production (*i.e.* "leading jet" events) and Drell-Yan lepton pair production in the mass region of the Z-boson (*i.e.* "Z-boson" events) have been presented and compared with several QCD Monte-Carlo model tunes. The data are corrected to the particle level and compared with the Monte-Carlo models at the particle level (*i.e.* generator level). The "underlying event" is similar for "leading jet" and "Z-boson" events as one would expect. The goal of the CDF analysis is to provide data that can be used to tune and improve the QCD Monte-Carlo models of the "underlying event" that are used to simulate hadron-hadron collisions. It is important to tune the new QCD Monte-Carlo MPI models [10, 11] so that we can begin to use them in data analysis. I believe once the new QCD Monte-Carlo models have been tuned that they will describe the data better than the old Pythia 6.2 tunes (see the talks by Peter Skands and Hendrik Hoeth as this conference).

PYTHIA Tune A and Tune AW do a good job in describing the CDF data on the "underlying event" observables for "leading jet" and "Z-boson" events, respectively, although the agreement between theory and data is not perfect. The "leading jet" data show slightly more activity in the "underlying event" than PYTHIA Tune A. PYTHIA Tune AW is essentially identical to Tune A for "leading jet" events. All the tunes with MPI agree better than HERWIG without MPI. This is especially true in the "toward" region in "Z-boson" production. Adding JIMMY MPI to HERWIG greatly improves the agreement with data, but HERWIG with JIMMY MPI produces a charged particle p_T spectra that is considerably "softer" than the data. The PYTHIA ATLAS tune also produces a charged particle p_T spectra that is considerably "softer" than the data.

The behavior of the average charged particle p_T versus the charged particle multiplicity is an important observable. The rate of change of $<p_T>$ versus charged multiplicity is a measure of the amount of hard versus soft processes contributing and it is sensitive the modeling of the multiple-parton interactions. PYTHIA Tune A and Tune AW do a good job in describing the data on $<p_T>$ versus multiplicity for "min-bias" and "Z-boson" events, respectively, although again the agreement between theory and data is not perfect. The behavior of $<p_T>$ versus multiplicity is remarkable similar for "min-bias" events and "Z-boson" events with $p_T(Z) < 10$ GeV/c suggesting that MPI are playing an important role in both these processes.

Models with multiple-parton interactions like PYTHIA Tune DWT predict that the "underlying event" will become much more active (with larger $<p_T>$) at the LHC. For HERWIG (without MPI) the "toward" region of "Z-boson" production does not change much in going from the Tevatron to the LHC. It is important to measure the "underlying event" observables presented here as soon as possible at the LHC. We will learn a lot about MPI by comparing the Tevatron results with the early LHC measurements.

References and Footnotes

[1] T. Sjostrand, Phys. Lett. **157B**, 321 (1985); M. Bengtsson, T. Sjostrand, and M. van Zijl, Z. Phys. **C32**, 67 (1986); T. Sjostrand and M. van Zijl, Phys. Rev. **D36**, 2019 (1987). T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna and E. Norrbin, Computer Physics Commun. **135**, 238 (2001). We use PYTHIA version 6.216.

[2] *Hard Underlying Event Corrections to Inclusive Jet Cross-Sections*, Jon Pumplin, Phys. Rev. **D57**, 5787 (1998).

[3] *Charged Jet Evolution and the Underlying Event in Proton-Antiproton Collisions at 1.8 TeV*, The CDF Collaboration (T. Affolder et al.), Phys. Rev. **D65**, 092002, (2002).

[4] *Measurement of the Z PT Distribution in Proton-Antiproton Collisions at 1.8 TeV*, The CDF Collaboration (F. Abe et al.), Phys. Rev. Lett. **67**, 2937-2941 (1991).

[5] The value of PARP(62), PARP(64), and PARP(91) was determined by CDF Electroweak Group. The "W" in Tune AW, BW, DW, DWT, QW stands for "Willis". I combined the "Willis" tune with Tune A, etc..

[6] Phys. Rev. Lett. **94**, 221801 (2005).

[7] J.M. Butterworth, J.R. Forshaw, and M.H. Seymour, Z. Phys. **C7**, 637-646 (1996).

[8] G. Marchesini and B. R. Webber, Nucl. Phys **B310**, 461 (1988); I. G. Knowles, Nucl. Phys. B310, 571 (1988); S. Catani, G. Marchesini, and B. R. Webber, Nucl. Phys. **B349**, 635 (1991).

[9] *Measurement of Particle Production and Inclusive Differential Cross Sections in Proton-Antiproton Collisions at 1.96 TeV*, N. Moggi et. al. (CDF Collaboration), CDF/PUB/MIN BIAS/PUBLIC/9550, *to be submitted to Phys. Rev. D*.

[10] T. Sjostrand and P. Z. Skands, Eur. Phys. J., **C39** 129, (2005). T. Sjostrand, S. Mrenna and P. Skands, JHEP 05 (2006) 026.

[11] P. Skands and D. Wicke, Eur. Phys. J. **C52** 133, (2007).

Recent Progress in Jet Algorithms and Their Impact in Underlying Event Studies

Matteo Cacciari [1,2]
[1]LPTHE, UPMC – Paris 6, CNRS UMR 7589, Paris, France
[2]Université Paris-Diderot – Paris 7, Paris, France

Abstract
Recent developments in jet clustering are reviewed. We present a list of fast and infrared and collinear safe algorithms, and also describe new tools like jet areas. We show how these techniques can be applied to the study of underlying event or, more generally, of any background which can be considered distributed in a sufficiently uniform way.

1 Recent Developments in Jet Clustering

The final state of a high energy hadronic collision is inherently extremely complicated. Hundreds or even thousand of light hadrons and leptons can be recorded by modern detectors, making the task of reconstructing the original (simpler) hard event very difficult. This large number of particles is the product of a number of branchings and decays which follow the initial production of a handful of partons. Usually only a limited number of stages of this production process can be meaningfully described in quantitative terms, for instance by perturbation theory in QCD. This is why, in order to compare theory and data, the latter must first be *simplified* down to the level described by the theory.

Jet definitions offer precisely this possibility of creating calculable observables from many final-state particles. This is done by clustering them into jets via a well specified algorithm, which usually contains one or more parameters, the most important of them being a "radius" R which controls the extension of the jet in the rapidity-azimuth plane. One can also choose a recombination scheme, which controls how partons' (or jets') four-momenta are combined. The combination of a *jet algorithm*, its *parameters* and the *recombination scheme* is called a *jet definition* [1], and must be specified in full (together with the initial particles sample) in order for the process

$$\{\text{particles}\} \overset{\text{jet definition}}{\longrightarrow} \{\text{jets}\} \tag{1}$$

to be fully reproducible and the final jets to be the same.

While (almost) any jet definition can produce sensible observables, not all of them will produce one which is *calculable* in perturbation theory. For the latter to be true, the jet algorithm must be *infrared and collinear safe* (IRC safe) [2], meaning that actions producing configurations that lead to divergences in perturbation theory, namely the emission of a soft particle or a collinear splitting of a particle into two) must not produce any change in the jets returned by the algorithm.

The importance for jet algorithms to be IRC safe had been recognized as early as 1990 in the 'Snowmass accord' [3], together with the need for them to be easily applicable both on the theoretical and the experimental side. However, many of the implementations of jet clustering

Jet algorithm	Type of algorithm, (distance measure)	algorithmic complexity
k_t [5, 6]	SR, $d_{ij} = \min(k_{ti}^2, k_{tj}^2)\Delta R_{ij}^2/R^2$	$N \ln N$
Cambridge/Aachen [7, 8]	SR, $d_{ij} = \Delta R_{ij}^2/R^2$	$N \ln N$
anti-k_t [10]	SR, $d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2})\Delta R_{ij}^2/R^2$	$N^{3/2}$
SISCone [9]	seedless iterative cone with split-merge	$N^2 \ln N$

Table 1: List of some of the IRC safe algorithms available in `FastJet`. SR stands for 'sequential recombination'. k_{ti} is a transverse momentum, and the angular distance is given by $\Delta R_{ij}^2 = \Delta y_{ij}^2 + \Delta\phi_{ij}^2$.

algorithms used in the following decade and a half failed to provide these characteristics: cone-type algorithms were typically infrared or collinear unsafe beyond the two or three particle level (see [1] for a review), whereas recombination-type algorithms were usually considered too slow to be usable at the experimental level in hadronic collisions.

This deadlock was finally broken by two papers, one in in 2005 [4], which made sequential recombination type clustering algorithms like k_t [5, 6] and Cambridge/Aachen [7, 8] fast, and one in 2007, which introduced SISCone [9], a cone-type algorithm which is infrared and collinear safe. A further paper introduced in 2008 the anti-k_t algorithm [10], a fast, IRC safe recombination-type algorithm which however behaves, for many practical purposes, like a nearly-perfect cone. This set of algorithms (see Table 1), all available through the `FastJet` package [11], allows one to replace most of the unsafe algorithms still in use with fast and IRC safe ones, while retaining their main characteristics (for instance, the MidPoint and the ATLAS cone could be replaced by SISCone, and the CMS cone could be replaced by anti-k_t).

2 Jet Areas

A by-product of the speed and the infrared safety of the new algorithms (or new implementations of older algorithms) was found to be the possibility to define in a practical way the *area* of a jet, which measures its susceptibility to be contaminated by a uniformly distributed background of soft particles in a given event.

In their most modest incarnation, jet areas can be used to visualize the outline of the jets returned by an algorithm so as to appreciate, for instance, if it returns regular ("conical") jets or rather ragged ones. An example is given in Fig. 1.

Jet areas are amenable, to some extent, to analytic treatments [13], or can be measured numerically with the tools provided by `FastJet`. These analyses disprove the common assumption that all cone-type algorithms have areas equal to πR^2. In fact, depending on exactly which type of cone algorithm one considers, its areas con differ, even substantially so, from this naive estimate: for instance, the area of a SISCone jet made of a single hard particle immersed in a background of many soft particles is rather $\pi R^2/4$. This little catchment area can explain why iterative cone algorithms with a split-merge procedure (like the MidPoint algorithm in use at CDF) have often been seen to fare 'well' in noisy environments. One can analyse next the k_t and the Cambridge/Aachen algorithms, and see that their single-hard-particle areas turn out to be roughly $0.81\pi R^2$. Finally, this area for the anti-k_t algorithm is instead exactly πR^2. This fact,

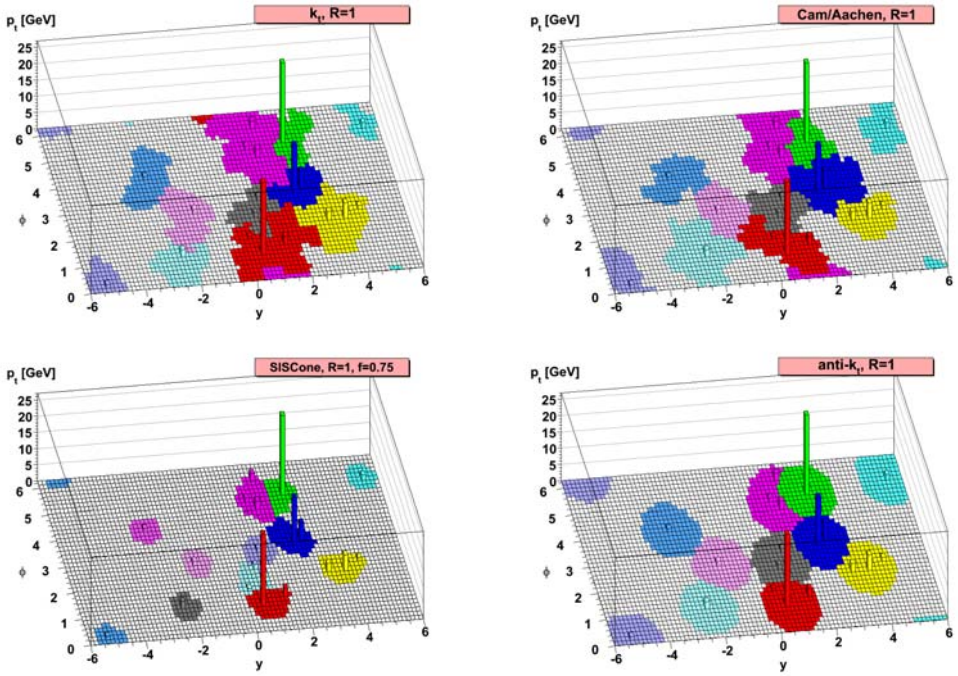

Fig. 1: Typical jet outlines returned by four different IRC safe jet clustering algorithms. From [10].

together with its regular contour shown in Fig. 1, explains why it is usually considered to behave like a 'perfect cone'.

Jet areas also allow one to use some jet algorithms as tools to measure the level of a sufficiently uniform background which accompanies the harder events. This can be accomplished by following the procedure outlined in [14]: for each event, all particles are clustered into jets using either the k_t or the Cambridge/Aachen algorithms, and the transverse momentum $p_{t,j}$ and the area A_j of each jet are calculated. One observes that a few hard jets have large values of transverse momentum divided by area, whereas most of the other, softer jets have similar (and smaller) values of this ratio. The background level ρ, transverse momentum per unit area in the rapidity-azimuth plane, is then obtained as

$$\rho = \text{median} \left\{ \frac{p_{t,j}}{A_j} \right\}_{j \in \mathcal{R}} . \tag{2}$$

The range $\mathcal{R}$ should be the largest possible region of the rapidity-azimuth plane over which the background is expected to be constant.

The operation of taking the median of the $\{p_{t,jet}/A_{jet}\}$ distribution is, to some extent, arbitrary. It has been found to give sensible results, provided that the range $\mathcal{R}$ contains sufficiently many soft background jets – at least about ten (twenty) of them, if only one (two) harder jets are also present, are usually enough.

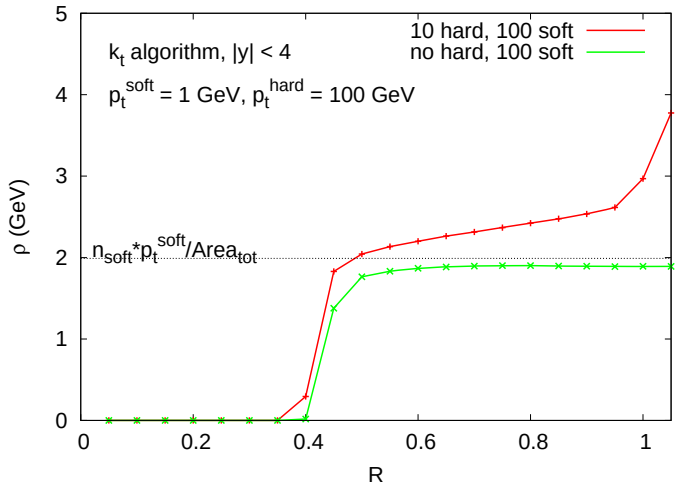

Fig. 2: Determination of the background level ρ of a toy-model random underlying event, as a function of the radius parameter R. Each point is the result of averaging over many different realizations. The parameters have been adjusted to roughly reproduce the situation expected at the LHC.

3 Underlying Event Studies

To a certain extent, and within certain limits, the background to a hard collision created by the soft particles of the underlying event (EU) can be considered fairly uniform. It becomes then amenable to be studied with the technique introduced in the previous Section. This constitutes an alternative to the usual and widespread approach of triggering on a leading jet, and selecting the two regions in the azimuth space which are transverse to its direction and that of the recoil jet. These two regions are considered to be little affected by hard radiation (in the least energetic of them it is expected to be suppressed by at least two powers of α_s), and therefore one can expect to be able to measure the UE level there.

This way of selecting the UE can be considered a *topological* one: particles (or jets) are classified as belonging to the UE or not as a result of their position. On the other hand, the median procedure described in the previous Section can be thought of as a *dynamical selection*: no a priori hypotheses are made and, in a way that changes from one event to another, a jet is automatically classified as belonging to the hard event or to the background as a result of its characteristics (namely the value of the $p_{t,j}/A_j$ ratio). One can further show that this selection pushes the possible contamination from perturbative radiation to very large powers of α_s: for a range $\mathcal{R}$ defined by $|y| < y_{max}$, perturbative contamination will only start at order $n \simeq 3y_{max}/R^2$ [14]. This gives $n \sim 24$ for $y_{max} = 2$ and $R = 0.5$, suggesting that the perturbative contribution is minimal.

A sensible criticism of this procedure is that the UE distribution is not necessarily uniform,

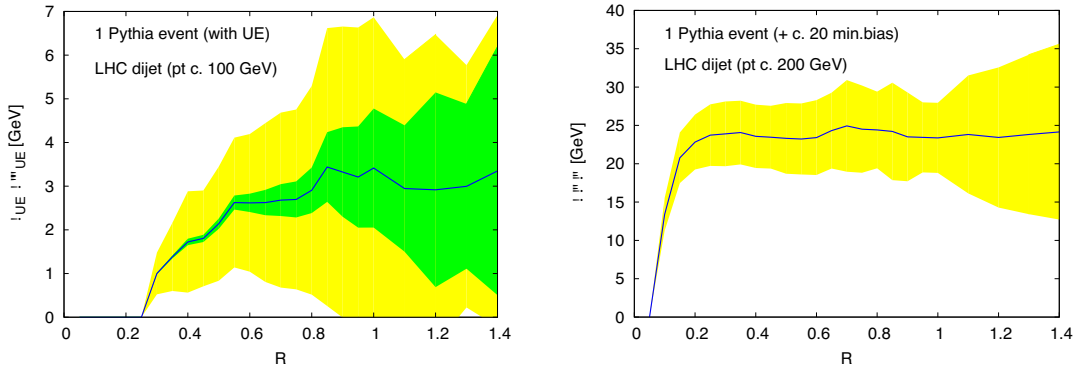

Fig. 3: Determination of the background level ρ in realistic dijet events at the LHC, with (right) and without (left) pileup. Preliminary results.

and may for instance vary as a function of rapidity. A way around this is then to choose smaller ranges, located at different rapidity values, and repeat the ρ determination in each of them. Of course care will have to be taken that the chosen ranges remain large enough to satisfy the criterion on the number of soft jets versus hard ones given in the previous Section: for instance, a range one unit of rapidity large can be expected to contain roughly $2\pi/(0.55\pi R^2) \sim 15$ soft jets for $R = 0.5$, which makes it marginally apt to the task[1].

A final word should be spent on which values of the radius parameter R can be considered appropriate for this analysis. Roughly speaking, R should be large enough for the number of 'real' jets (i.e. containing real particles) to be at last larger than the number of 'empty jets' (regions of the rapidity-azimuth plane void of particles, and not occupied by any 'real' jet). It should also be small enough to avoid having too many jets containing too many hard particles. Analytical estimates [14] and empirical evidence show that for UE estimation in typical LHC conditions one can expect values of the order of 0.5 – 0.6 to be appropriate. Much smaller values will return $\rho \simeq 0$, while larger values will tend to return progressively larger values of ρ, as a result of the increasing contamination from the hard jets. Fig. 2 shows results obtained with a toy model where 100 soft particles with $p_T^{soft} \simeq 1$ GeV are generated in a $|y| < 4$ region. Ten hard particles, with $p_T^{hard} \simeq 100$ GeV, can be additionally generated in the same region. One observes how, after a threshold value for R, ρ is estimated correctly for the soft-only case, while when hard particles are present they increasingly contaminate the estimate of the background.

The same analysis can be performed on more realistic events, generated by Monte Carlo simulations. Fig. 3 shows the determination of ρ in a simulated dijet event at the LHC, with and without pileup. In both cases the general structure of the toy-model in Fig. 2 can be seen, though it is worth noting that in the UE case (left plot) the slope can vary significantly from event to event, and also according to the Monte Carlo tune used [15]. The larger particle density (and probably higher uniformity) of the pileup case allows for an easier and more stable determination.

Once a procedure for determining ρ is available, one can think of many different appli-

[1] Its performance can be improved by removing the hardest jets it contains from the $\{p_{t,j}/A_j\}$ list before taking the median [15].

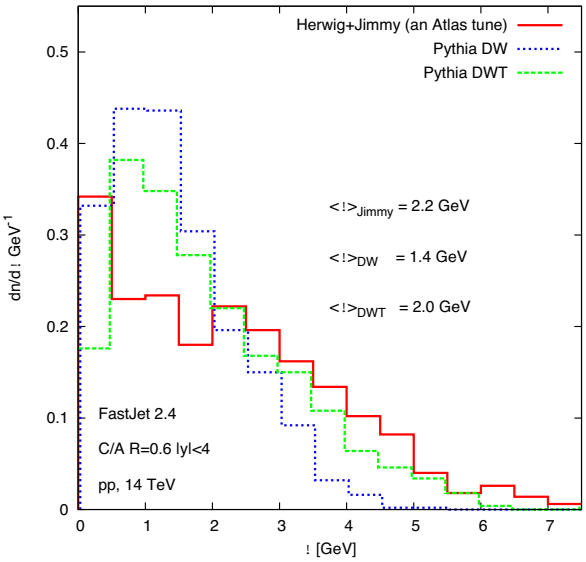

Fig. 4: Distributions of ρ from the UE over many simulated LHC dijet events ($p_T > 50$ GeV, $|y| < 4$), using different Monte Carlos and different UE tunes. Preliminary results.

cations. One possibility is of course to tune Monte Carlo models to real data by comparing rho distributions, correlations, etc. A preliminary example is given in fig. 4, where studying the distribution of ρ can be seen to allow one to discriminate between UE models which would otherwise give similar values for the average contribution $\langle\rho\rangle$. More extensive studies are in progress [15].

Yet another use of measured ρ values is the *subtraction* of the background from the transverse momentum of hard jets. Ref. [14] proposed to correct the four-momentum $p_{\mu j}$ of the jet j by an amount proportional to ρ and to the area of the jet itself (the susceptibility of the jet to contamination):

$$p_{\mu j}^{sub} = p_{\mu j} - \rho A_{\mu j} \tag{3}$$

where $A_{\mu j}$ is a four-dimensional generalization of the concept of jet area, normalized in such a way that its transverse component coincides, for small jets, with the scalar area A_j [13]. One can show [14, 16] that such subtraction of the underlying event can improve in a non-negligible way the reconstruction of mass peaks even at very large energy scales. A similar procedure is also being considered [17] for heavy ion collisions, where the background can contribute a contamination even larger than the transverse momentum of the hard jet itself (partly because of this, one usually talks of 'jet reconstruction' in this context, rather than just 'subtraction'). Initial versions of this technique have already been employed at the experimental level by the STAR Collaboration at RHIC in [18, 19], where IRC safe jets have been reconstructed for the first time

in heavy ion collisions.

4 Conclusions

Since 2005 numerous developments have intervened in jet physics. A number of fast and infrared and collinear safe algorithms are now available, allowing for great flexibility in analyses. Tools have been developed and practically implemented to calculate jet areas, and these can used to study various types of backgrounds (underlying event, pileup, heavy ions background) and also to subtract their contribution to large transverse-momentum jets.

These new algorithms and methods (as well as the ones not mentioned in this talk, like the many approaches to jet substructure, see e.g. [20–23], useful in a number of new-physics searches) are transforming jet physics from being just a way to obtain calculable observable to providing a full array of precision tools with which to probe efficiently the complex final states of high energy collisions.

Acknowledgments

I wish to thank the organizers of MPI@LHC'08 in Perugia for the invitation to this interesting conference, as well as Gavin P. Salam, Gregory Soyez, Juan Rojo and Sebastian Sapeta for the stimulating ongoing collaboration on jet issues.

References

[1] C. Buttar *et al.* (2008). 0803.0678 [hep-ph].

[2] G. Sterman and S. Weinberg, Phys. Rev. Lett. **39**, 1436 (1977).

[3] J. E. Huth *et al.* Presented at Summer Study on High Energy Physics, Reaearch Directions for the Decade, Snowmass, CO, Jun 25 - Jul 13, 1990.

[4] M. Cacciari and G. P. Salam, Phys. Lett. **B641**, 57 (2006). hep-ph/0512210.

[5] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, Nucl. Phys. **B406**, 187 (1993).

[6] S. D. Ellis and D. E. Soper, Phys. Rev. **D48**, 3160 (1993). hep-ph/9305266.

[7] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, JHEP **08**, 001 (1997). hep-ph/9707323.

[8] M. Wobisch and T. Wengler (1998). hep-ph/9907280.

[9] G. P. Salam and G. Soyez, JHEP **05**, 086 (2007). 0704.0292 [hep-ph].

[10] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 063 (2008). 0802.1189.

[11] M. Cacciari, G. P. Salam, and G. Soyez, http://www.fastjet.fr/.

[12] S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock, and B. R. Webber, Phys. Lett. **B269**, 432 (1991).

[13] M. Cacciari, G. P. Salam, and G. Soyez, JHEP **04**, 005 (2008). 0802.1188.

[14] M. Cacciari and G. P. Salam, Phys. Lett. **B659**, 119 (2008). 0707.1378.

[15] M. Cacciari, G. P. Salam, and S. Sapeta, in preparation.

[16] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, JHEP **12**, 032 (2008). 0810.1304.

[17] M. Cacciari, J. Rojo, G. P. Salam, and G. Soyez, in preparation.

[18] STAR Collaboration, S. Salur (2008). 0809.1609.

[19] STAR Collaboration, S. Salur (2008). 0810.0500.

[20] J. M. Butterworth, A. R. Davison, M. Rubin, and G. P. Salam, Phys. Rev. Lett. **100**, 242001 (2008). 0802.2470.

[21] J. Thaler and L.-T. Wang, JHEP **07**, 092 (2008). 0806.0023.

[22] D. E. Kaplan, K. Rehermann, M. D. Schwartz, and B. Tweedie, Phys. Rev. Lett. **101**, 142001 (2008). 0806.0848.

[23] L. G. Almeida *et al.* (2008). 0807.0234.

Part II

Soft and Hard Multiple Parton Interactions

Convenors:

Arthur Moraes (University of Glasgow)
Richard Field (Florida University)
Mark Strikman (Penn State University)

Soft and Hard Multiple Parton Interactions

Paolo Bartalini
National Taiwan University

In the years '80, the evidence for Double Scattering (DS) phenomena in the high-p_T phenomenology of hadron colliders suggests the extension of the same perturbative picture to the soft regime, giving rise to the first implementation of the Multiple Parton Interaction (MPI) processes in a QCD Monte Carlo model by T.Sj *o*strand and M.van Zijl. Such model turns out to be very successful in reproducing the UA5 charged multiplicity distributions and in accounting for the violation of the sensitive Koba Nielsen Olesen scaling violation at increasing center of mass energies.

The implementation of the MPI in the QCD Monte Carlo models is quickly proceeding through an increasing level of sophistication and complexity, still leaving room for different approaches and further improvements like the introduction of a dynamical quantum description of the interacting hadrons providing a modeling of the diffractive interactions in the same context. See the detailed discussion in the introduction of Section IV.

As deeply discussed both in Section I and Section II, considerable progress in the phenomenological study of the Underlying Event (UE) in jet events is achieved by the CDF experiment at the Tevatron collider, with a variety of redundant measurements relying both on charged tracks and calorimetric clusters, the former being intrinsically free from the pile-up effects and achieving a better sensitivity at low p_T. Challenging tests to the universality features of the models are provided by the extension of the UE measurement to the Drell Yan topologies and by the additional complementary measurements on MB events dealing with the correlations between charged multiplicity and average charged momentum.

While preparing the ground for the traditional Minimum Bias (MB),Underlying Event (UE) and Double Scattering (DS) measurements at the LHC along the precious Tevatron experience also complemented with the recent UE HERA results, new feasibility studies are proposed which in perspective will constitute a challenge to the predictivity and to the consistency of the models: the usage of jet clustering algorithms providing an automated estimation of the UE activity, the measurement of large pseudo-rapidity activity correlations, the investigation of the mini-jet structure of the MB events, the evaluation of the impact of the MPI on the total cross section.

With the LHC data taking period approaching, the experiments put a lot of emphasis on the physics validation and tuning of the models, in particular for what concerns the energy dependency of the parameters. The tune of the MPI parameters is a very delicate issue which has impact on the calibration of major physics tools like the vertex reconstruction and the isolation techniques.

A significant fraction of the early measurements of ALICE, ATLAS, CMS, LHCb and TOTEM will be affected by the MPI, with most of the LHC feasibility studies shown in these proceedings turned into physics publications in a reasonably short time scale. In other words the MPI will be one of the first features of the LHC physics which will be deeply tested with an high degree of complementarity and redundancy, and we should be ready for possible surprises!

Multiple Production of W Bosons in pp and pA Collisions

E. Braidot, E. Cattaruzza, A. Taracchini, D. Treleani[†]
Department of Theoretical Physics, University of Trieste and INFN, Section of Trieste

Abstract
The production of equal sign W boson pairs, through single and double parton collisions, are comparable in magnitude at the LHC. As a consequence of the strong anti-shadowing of MPI in interactions with nuclei, the double scattering contribution is further enhanced in the case of hadron-nucleus collisions

1 Multiple production of W bosons in proton-proton collisions

Multiple parton interactions are a manifestation of the unitarity problem caused by the rapid increase of the parton flux at small x, which leads to a dramatic growth of all cross sections with large momentum transfer in pp collisions at the LHC [5]. The critical kinematical regime may be identified by comparing the rate of double collisions with the rate of single collisions. When the two rates become comparable multiple collisions are no more a small perturbation and all multiple collisions become equally important, while the production of large p_t partons becomes a common feature of the inelastic event [10] [3]. In its simplest implementation [9] the double parton scattering cross section σ_D is given by

$$\sigma_D = \frac{1}{2}\frac{\sigma_S^2}{\sigma_{eff}} \tag{1}$$

where σ_S is the single scattering cross section. The problem with unitarity becomes hence critical in the kinematical domain where σ_S and the scale factor σ_{eff} are of the same order.

The experimental indication is that the value of σ_{eff} is close to 10 mb [1]. One might hence conclude that one should worry about multiple parton collisions only when the single scattering cross section becomes comparable with σ_{eff}. On the contrary multiple parton collisions may represent an important effect also in cases where the single scattering cross section is many orders of magnitude smaller that σ_{eff}. The consideration applies to the interesting case of the production of equal sign W boson pairs. The leptonic decay channel of W bosons, which leads to final states with isolated leptons plus missing energy, is in fact of great interest for the search of new physics [2].

The production of two equal sign W bosons is a higher order process in the Standard Model and two equal sign W bosons can be produced only in association with two jets [7]. At the lowest order there are 68 diagrams at $\mathcal{O}(\alpha_W^4)$ and 16 diagrams at $\mathcal{O}(\alpha_S^2\alpha_W^2)$ (some of the diagrams are shown in Fig.1) and, even though $\alpha_S > \alpha_W$, the strong and electroweak diagrams give comparable contributions to the cross section, which is infrared and collinear safe and can be evaluated without imposing any cutoff in the final state quark jets.

[†] speaker

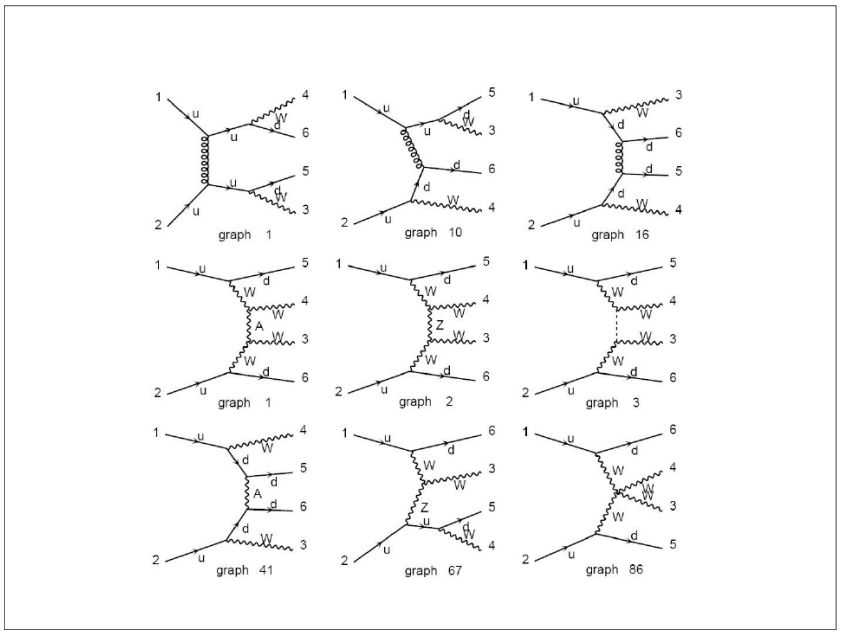

Fig. 1: Some of the three level diagrams which contribute to equal sign W pairs production

The resulting cross sections to produce W bosons and W boson pairs, by single parton scattering in pp interactions, are shown in Fig.2 as a function of the c.m. energy. As apparent in the figure (left upper panel) the cross section to produce two equal sign W bosons is five orders of magnitude smaller with respect to the cross section to produce a single W boson. The same reduction factor is expected for the production of two equal sign W bosons through a multiple collisions processes:

$$\sigma_{WW} = \frac{1}{2}\sigma_W \frac{\sigma_W}{\sigma_{eff}}, \qquad \frac{\sigma_W}{\sigma_{eff}} \simeq \frac{10^2 \mathrm{nb}}{10 \mathrm{mb}} = 10^{-5} \tag{2}$$

The argument above relies on the simplest expression of the double scattering cross section, obtained by assuming a factorized expression for the the double parton distributions, which is obviously inconsistent in the case of the valence because of the correlations induced by flavor conservation. In the actual case, given the large mass of the W bosons, one may expect important contributions of the valence also at the LHC. One may hence normalize the double parton distributions in such a way to satisfy the flavor sum rules and work out the double scattering cross section accordingly. The effect on the cross section is shown in the left lower panel of Fig.2, which shows that, at the LHC, the cross sections is reduced by about 20%.

The integrated rates of equal sign W boson pairs, by single and double parton collisions, are hence comparable in pp collisions at the LHC. The distribution in phase space is however rather different in the two cases.

In the right lower panel of Fig.2 we show the distribution of the produced Ws, as a function of their transverse momenta. The distribution in transverse momenta of the produced Ws is

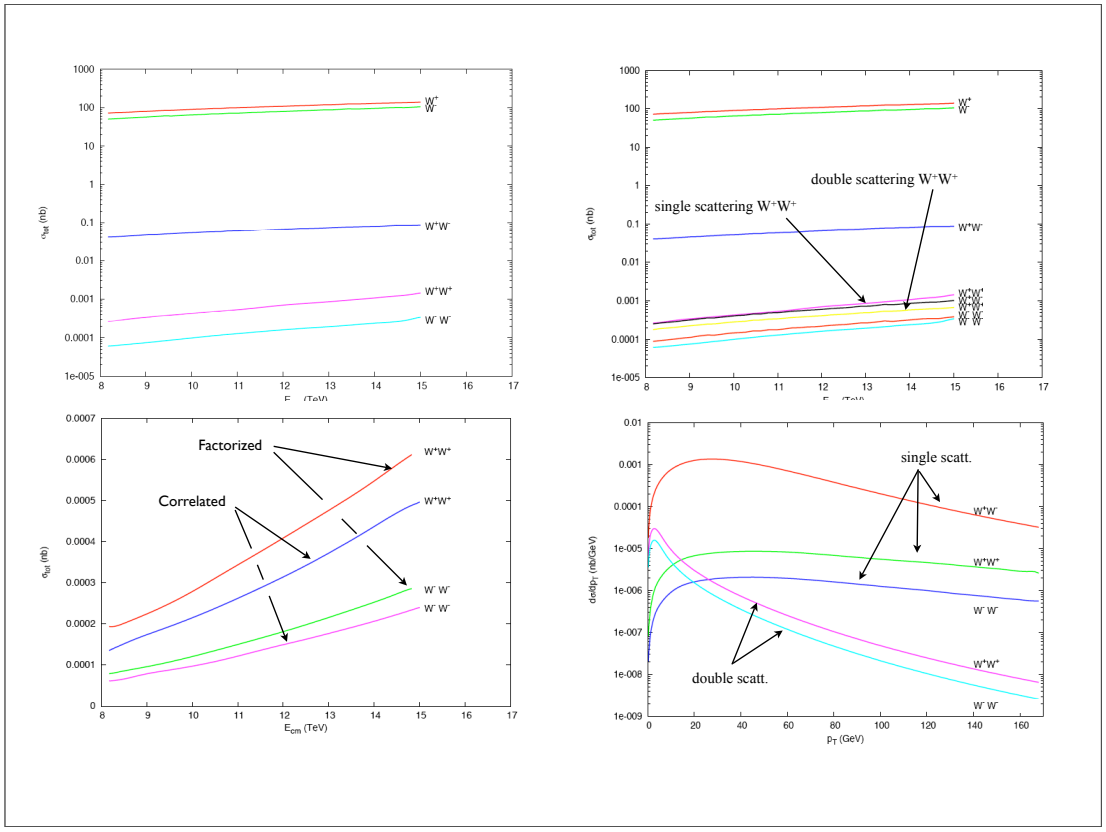

Fig. 2: *Upper left panel*: W production cross sections by single parton scattering in pp interactions as a function of the c.m. energy. *Upper right panel*: W and W pairs production cross sections in pp interactions by double and by single parton collisions. *Lower left panel*: W pairs production cross sections by double parton collisions with correlated and uncorrelated parton densities in the case of pp interactions. *Lower right panel*: W pairs densities in transverse space in the case of single and of double parton collisions in pp interactions.

obtained by following the recipe of the "Poor Man's shower model" of Barger and Phillips [4] and using as a smearing function at low p_t the expression in Eq.15 of [8]. The two contributions may be separated with a cut of 15 GeV/c in the transverse momenta of the produced Ws. In Fig. 3 we show how the W^+ bosons (left panels) and their decay electrons (right panels) are distributed in transverse momentum and rapidity. The case of double parton collisions is shown in the upper panels, while the case of single parton collisions is shown in the lower panels. In the case of a double parton collision, the W bosons are mainly produced with small transverse momenta, while the rapidity distribution of the W boson reminds the momentum of the originating up quarks. The distributions of the final state charged leptons is peaked at the same rapidity of the parent W boson and at a transverse momentum corresponding to 1/2 of the W boson mass.

In the case of single parton collisions (lower panels of Fig.3) the Ws and the corresponding decay leptons have a much broader distribution in p_t and rapidity and the characteristic peaks of the double scatterings are completely absent. The two contributions are hence disentangled very easily by adopting appropriate cuts in rapidity and transverse momenta of the finally observed charged leptons.

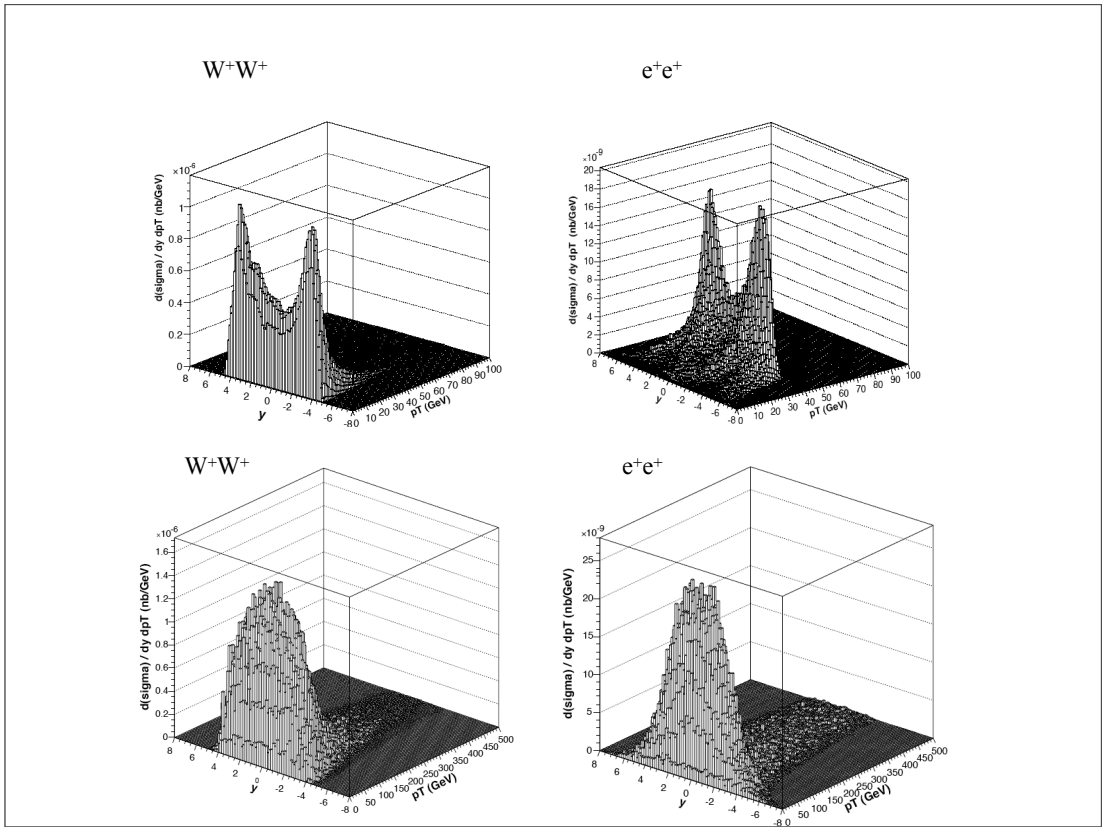

Fig. 3: W^+W^+ and e^+e^+ pairs distribution in transverse momentum and rapidity, in the case of single parton collision (upper panels) and of double parton collisions (lower panels) in proton-proton collisions.

2 Multiple production of W bosons in proton-nucleus collisions

As pointed out in [11], a major feature of MPI in hadron-nucleus collisions is the strong anti-shadowing. Double parton collisions may in fact be amplified by a factor 2 or 3 on heavy nuclei as compared with the corresponding cross section in hadron-nucleon collisions multiplied by the atomic mass number A. Notice that for, say, values of x of the order of 10^{-3} and for values of $Q^2 > 10$ GeV^2, the usual nuclear shadowing correction is a much smaller effect and corresponds to a reduction of the cross section not larger than 10% even on heavy nuclei [6]. The effect is schematically illustrated in Fig. 4, where non additive corrections to the nuclear structure functions are neglected, in such a way that each nuclear parton may be associated to a given parent nucleon. As shown in Fig.4, in proton-nucleus interactions one may hence distinguish two different contributions to the double parton scattering cross section, depending wether the two nuclear partons undergoing the interactions are originated by one or by two different target nucleons.

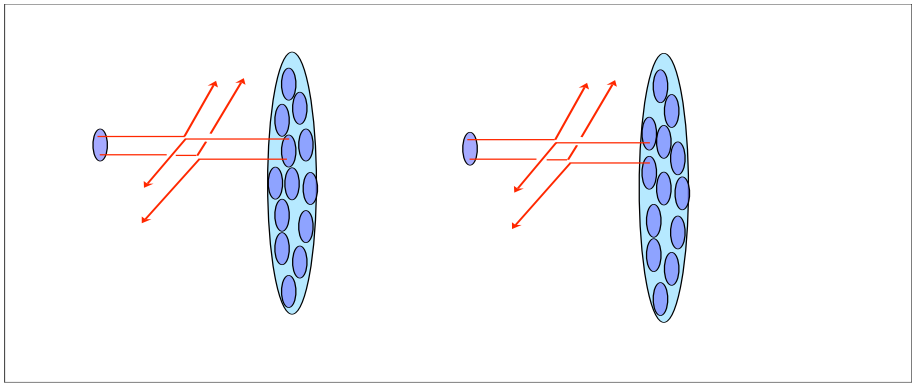

Fig. 4: W production cross sections by single parton scattering in pp collisions as a function of the c.m. energy.

The cross section may thus be written as the sum of two terms

$$\sigma_D^A = \sigma_D^A|_1 + \sigma_D^A|_2 \tag{3}$$

and

$$\sigma_D^A|_1 = \frac{1}{2}\frac{\sigma_W^2}{\sigma_{eff}} \int d^2bT(b) \propto A, \qquad \sigma_D^A|_2 = \frac{1}{2}\sigma_W^2 \int d^2bT^2(b) \propto A^{4/3}$$

The anti-shadowing effect is apparent in Fig.5, where the W production cross sections in proton-proton collisions are compared with the cross sections in proton-nucleus collisions (after dividing by the atomic mass number A). In the upper panels one compares the cross sections as a function of the c.m. energy, while in the lower panels one compares the distributions in transverse momenta of the two W^+ bosons. The region where double parton collisions dominate now extends to transverse momenta of the order of 40 GeV/c.

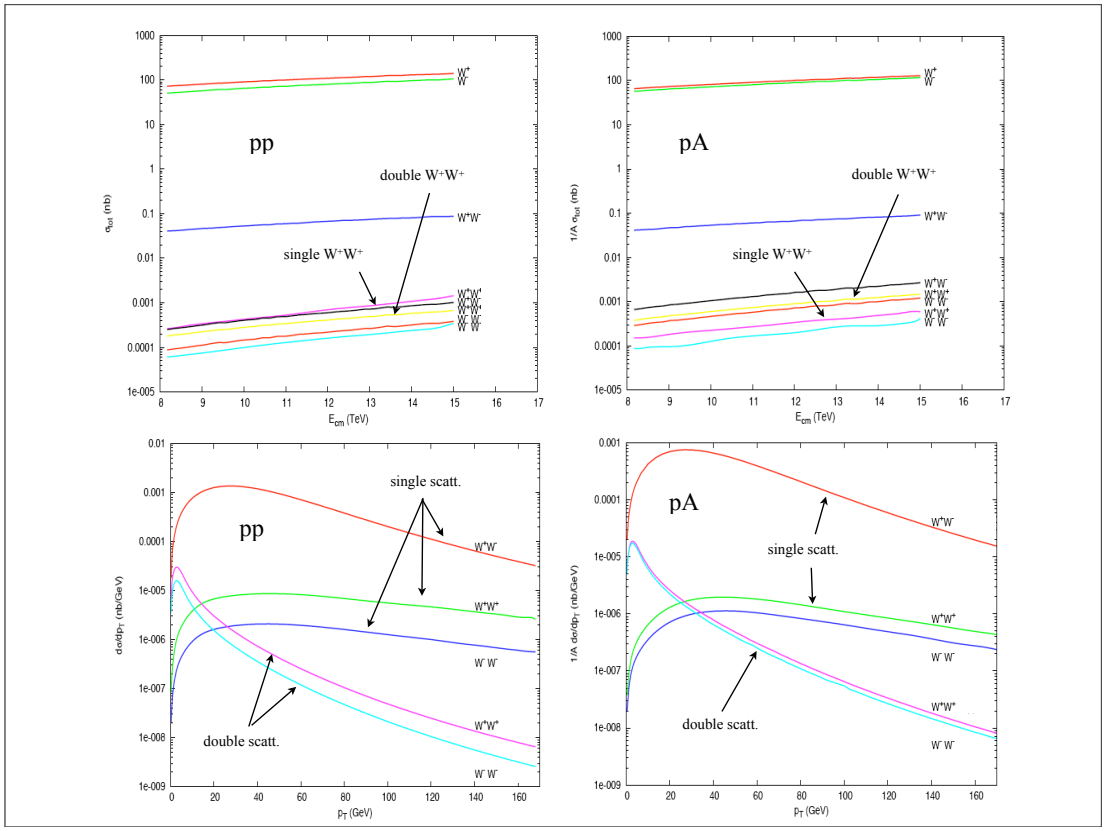

Fig. 5: W and W pairs production in proton-proton and proton-nucleus collisions. Integrated cross sections as a function of the c.m. energy (upper panels) and distributions in transverse space (lower panels).

In the upper panels of Fig.6 (left and right respectively) we show the distributions in transverse momentum and rapidity of the W^+ bosons and of the decay leptons in pA collisions. The W bosons are produced with a small transverse momentum, while the rapidity distribution of the W boson reminds the momentum of the originating up quark. The asymmetry in rapidity is due to the different content of up quarks in the proton as compared with the content of up quarks in the pairs of nucleons of the target nucleus undergoing the process (pp, pn and nn). The distributions of the final charged leptons is peaked at the same rapidity of the parent W boson and, as in the case of proton-proton interactions, at a transverse momentum corresponding to 1/2 of the W boson mass.

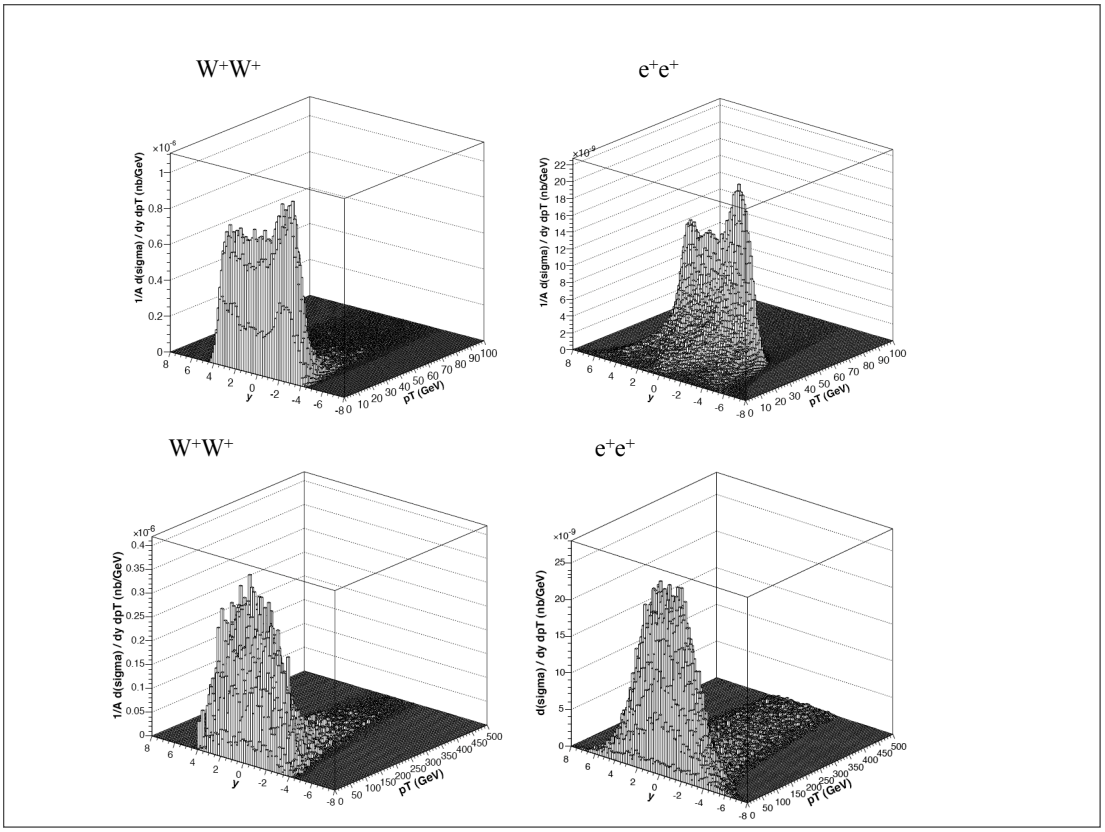

Fig. 6: W^+W^+ and e^+e^+ pairs distribution in transverse momentum and rapidity, in the case of single parton collision (upper panels) and of double parton collisions (lower panels) in proton-nucleus collisions.

The distributions of equal sign W bosons and of the decay leptons generated by single parton collisions in pA interactions are shown in the lower panels of Fig.6 (left and right respectively) as a function of rapidity and transverse momenta. The contribution of double collisions is overwhelming when selecting leptons with transverse momenta of the order of one half of the W mass.

3 Concluding summary

Equal sign W boson pairs are produced by a higher order process in the SM. As a consequence, the cross section to produce two W bosons with equal sign is more than two orders of magnitude smaller in pp collisions at the LHC, as compared with the cross section to produce two W bosons with opposite sign. An outcome is that the integrated cross sections, to produce two equal sign W bosons through single and double parton collisions, are similar in magnitude. The equal sign W bosons and the corresponding decay leptons are however distributed very differently in phase space by the two production mechanisms, which allows to disentangle the two contributions easily by looking at the distribution of the decay leptons.

As a consequence of the strong anti-shadowing of MPI in collisions with nuclei, the contribution of double scattering is greatly enhanced in the case of hadron-nucleus collisions.

References

[1] F. Abe et al. Double parton scattering in $\bar{p}p$ collisions at $\sqrt{s} = 1.8$TeV. *Phys. Rev.*, D56:3811–3832, 1997.

[2] A. Abulencia et al. Inclusive search for new physics with like-sign dilepton events in $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96-TeV. *Phys. Rev. Lett.*, 98:221803, 2007.

[3] L. Ametller and D. Treleani. SHADOWING IN SEMIHARD INTERACTIONS. *Int. J. Mod. Phys.*, A3:521–530, 1988.

[4] Barger, Vernon D. and Phillips, R. J. N. COLLIDER PHYSICS. *REDWOOD CITY, USA: ADDISON-WESLEY (FRONTIERS IN PHYSICS, 71)*, 1987.

[5] S. Catani et al. QCD. 2000.

[6] S. A. Kulagin and R. Petti. Global study of nuclear structure functions. *Nucl. Phys.*, A765:126–187, 2006.

[7] Anna Kulesza and W. James Stirling. Like sign W boson production at the LHC as a probe of double parton scattering. *Phys. Lett.*, B475:168–175, 2000.

[8] Anna Kulesza and W. James Stirling. Soft gluon resummation in transverse momentum space for electroweak boson production at hadron colliders. *Eur. Phys. J.*, C20:349–356, 2001.

[9] N. Paver and D. Treleani. MULTI - QUARK SCATTERING AND LARGE P(T) JET PRODUCTION IN HADRONIC COLLISIONS. *Nuovo Cim.*, A70:215, 1982.

[10] Torbjorn Sjostrand and Maria van Zijl. A Multiple Interaction Model for the Event Structure in Hadron Collisions. *Phys. Rev.*, D36:2019, 1987.

[11] M. Strikman and D. Treleani. Measuring double parton distributions in nucleons at proton nucleus colliders. *Phys. Rev. Lett.*, 88:031801, 2002.

Aknowledgements

The Organizing Committees would like to thank all the authors for their very high quality contributions and their participation to this first edition of the *Multiple Parton Interactions at the LHC Workshop*, the $I.N.F.N.$-$Perugia$ and the *Physics Department of the Università degli Studi di Perugia* for their support, and the *Comune di Perugia* for its contribution, particularly in providing the conference halls. Special thanks go to Francesca Rossi for providing and setting up the *Auditorium S. Cecilia*.