Measurement of the Energy Flow at Large Pseudorapidities for $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV at the Large Hadron Collider using the Compact Muon Solenoid

Dissertation

zur Erlangung des Doktorgrades

des Department Physik

der Universität Hamburg

vorgelegt von

Niladri Sen

aus Birkenhead

Hamburg 2011

Gutachter/Gutachterin der Dissertation:	Prof. Dr. Peter Schleper Priv.-Doz. Dr. Hannes Jung
Gutachter/Gutachterin der Disputation:	Prof. Dr. Joachim Mnich Priv.-Doz. Dr. Hannes Jung
Datum der Disputation:	24.10.2011
Vorsitzende/Vorsitzender des Prüfungsausschusses:	Prof. Dr. Caren Hagner
Vorsitzender des Promotionsausschusses:	Prof. Dr. Peter Hauschildt
Dekanin/Dekan des Fachbereichs Physik:	Prof. Dr. Heinrich Graener

Abstract

Energy flow, $dE/d\eta$, has been measured in proton-proton collisions at the LHC, for two centre-of-mass energies, $\sqrt{s} = 0.9$ TeV and 7 TeV, using an integrated luminosity of 239 μb^{-1} and 206 μb^{-1} respectively. The measurements were made in a previously unexplored phase space ($3.15 < |\eta| < 4.9$) using the CMS detector for two separate event topologies: minimum bias events and events with a hard scale set by the transverse momentum of the jets in a di-jet system. Data from each of the measurements have been compared to leading order Monte Carlo pp-collision event generators that use k_T^2-, Q^2- and angular-ordered parton showers. The forward energy flow measurements are shown to be sensitive to the models and tuning parameters in both their shape and magnitude. The necessity of underlying event models in order to describe data will be demonstrated. In addition, predictions from cosmic-ray event generators are shown to describe data consistently well for each of the measurements.

Zusammenfassung

In dieser Arbeit wurde der Energiefluss $dE/d\eta$ in Proton-Proton-Kollisionen des Beschleunigers LHC für zwei verschiedene Schwerpunktsenergien, $\sqrt{s} =$ 0.9 TeV und 7 TeV, mit Hilfe einer integrierten Luminosität von 239 μb^{-1} und 206 μb^{-1} gemessen. Die Messungen wurden in einem zuvor unerforschten Bereich des Phasenraums ($3.15 < |\eta| < 4.9$) mit Hilfe des CMS Detektors für zwei verschiedene Ereignistopologien durchgeführt: zum Einen wurden Minimum Bias Ereignisse betrachtet und zum Anderen wurden Zwei-Jet-Ereignisse analysiert, wobei die harte Skala durch den Transversalimpuls der Jets definiert wurde. Die Daten dieser beiden Messungen wurden mit Monte Carlo *pp*-Ereignisgeneratoren in fuehrender Ordnung verglichen, wobei k_T^2-, Q^2- und nach Winkeln sortierte Parton Shower Verwendung fanden. Diese Analysen zeigten, dass die Verteilungen des Energieflusses im Vorwärtsbereich sensitiv auf die unterschiedlichen physikalischen Modelle und die Tuning Parameter sowohl in ihrer Form als auch in ihrer Größenordnung reagierten. Es wird gezeigt, dass die Verwendung passender Underlying Event Modelle zur Beschreibung der Daten notwendig sind. Ferner wird aufgezeigt, dass die Vorhersagen von Ereignisgeneratoren für kosmische Strahlung die Daten der einzelnen Messungen konsistent beschreiben.

Declaration

I herewith declare that I have produced this thesis without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This thesis has not previously been presented in identical or similar form to any other German or foreign examination board.

The thesis work was conducted from 3rd October 2008 to 30th September 2011 under the supervision of Priv.-Doz. Dr. Hannes Jung at Deutsches-Elektronen Synchrotron.

Signed: Date:

To my mother and my father,
Ranjini and Himadri Sen

Contents

List of Figures

List of Tables

1

Introduction

Particle physics, the branch of science that studies the existence and interactions of the fundamental constituents of matter, has undergone centuries of conceptual development and seen many philosophical treaties on the nature of elementary particles. Abstract reasoning led ancient philosophers such as Leucippus and Democritus to the concept of fundamental building blocks of matter, the doctrine of atomism [1]. These ideas were further expounded upon by medieval and modern European scientists, each of whom attempted to reconcile contemporary theories on motion and mass with their own hypotheses. By the 19^{th} century, however, the majority of the beliefs were either riddled with speculation or rife with dogma. It was not until the investigations into the disintegration of elements by Rutherford, that the initial postulates were substantiated through empirical observation and experimentation. With the advent of the Wilson Cloud Chamber, the foundations for further investigations were set, giving birth to the field of experimental particle physics [2].

Modern particle physics focuses on subatomic particles by investigating the interactions between matter and energy at short distances, or equivalently, at high energies. As many of these elementary particles do not occur under ambient conditions on Earth, they are created artificially in high energy collisions by particle accelerators. Accelerator-based scattering experiments in the 1950s and 1960s detected a plethora of particles; the large variety hinted at more fundamental constituents, just as Mendeleev's periodic table suggested an atomic substructure. In an attempt to explain this myriad of particles as combinations of a smaller number of fundamental objects, the Standard Model (SM) [3] was formulated.

1. INTRODUCTION

The Standard Model [4] [5] is a collection of quantum field theories which respect the principles set by Special Relativity [6]. Using a Lagrangian [7] to define the relationship between the potential and kinetic energy of a field, observable forces can be deduced from local gauge symmetries within the Lagrangian density. Heisenberg's uncertainty principle [8] allows for the treatment of localised particles as extended fields, while the Lorentz-invariance [9] permits matter-energy equivalence. In short, the paradigm fulfils all the conditions required from a description of relativistic objects at the quantum scale. Moreover, the model has withstood almost all experimental verification conducted till date. It has proven to be robust in its predictions, and tested to a high precision up to energies of the order of several hundred giga-electron volts ($\mathcal{O}(100\text{ GeV})$). This includes recent measurements of neutrino masses, which have been adapted into the model through a simple expansion of the existing framework [10].

The exception is the elusive Higgs boson [11] [12] [13] [14]; a scalar particle produced from the electroweak symmetry breaking mediator SU(2)-doublet (the Higgs-field), for which no conclusive experimental evidence has been found till date [10]. Additionally, from a theoretical standpoint, the SM has a few shortcomings. It fails to address some fundamental issues, such as the how and why of fermion masses and mixing, the existing matter-antimatter asymmetry, and the gauge hierarchy problem [15] [16] [17]. More fundamentally, SM does not attempt to explain gravitation and cannot account for the observed amount of cold dark matter (CDM) [18]. Within the emerging "standard model of cosmology", composite particles such as baryons only account for four percent of the total matter and energy in the universe.

There are theories that adapt the current framework and propose certain changes to provide solutions to these failings. Alternatives include theories such as dynamical electroweak symmetry breaking via Technicolor [19] [20] [21] [22] [23], "little Higgs solutions" [24], low-energy supersymmetry [25] [26] and extra dimensions [27] [28] [29] [30] [31]. The majority of such scenarios predict signals at the TeV-scale, which, until recently, was beyond the scope of most particle accelerators.

Built to study this new energy scale is the Large Hadron Collider (LHC), a proton-proton collider with a beam energy of up to 7 TeV and a design luminosity of $L = 10^{34}\text{ cm}^{-2}\text{ s}^{-1}$. As it can produce a centre-of-mass energy seven times greater and luminosity one-hundred times higher than the TEVATRON, it is perfect for studying rare, high energy processes. However, the use of a hadron collider provides complications of

its own; proton-proton collisions are complex processes due to their internal hadronic structure. Often viewed as a superposition of several contributions, a collision event can be separated into two components: the primary scattering, and all the activity that is not associated with the primary process or the underlying event. As higher than before centre-of-mass energies ($\sqrt{s}$) are explored, the impact of the underlying event in collision events becomes progressively important.

On a theoretical level, components of the underlying event are not well understood, and a systematic description in Quantum Chromodynamics (QCD) remains a challenge. Nonetheless, phenomenological approaches to multi-parton dynamics exist, relying heavily on parametrised models (tunes) to describe data. But these tunes are limited as their extrapolation to larger energies is uncertain, often leading to sizeable discrepancies amongst the various predictions. The data collected with the Compact Muon Solenoid detector (CMS) allow for these models to be studied with far greater precision over a larger phase space than was possible in previous investigations.

Till date, measurements of the underlying event structure have been performed for central pseudorapidities (η). This thesis describes the measurement of the energy flow in the pseudorapidity range $3.15 < |\eta| < 4.9$. Two different centre-of-mass energies, 0.9 TeV and 7 TeV, are investigated, providing an insight into the scaling of the energy flow with respect to both the QCD scale and the centre-of-mass energy. A measurement of the energy flow in this η-region has never been reported at hadron colliders. Comparable measurements were performed in deep inelastic scattering events at HERA [32].

Aspects of QCD that are relevant to this measurement are summarised in Chapter 2. Chapter 3 describes the experimental apparatus. The main features of the Monte Carlo event simulations used for describing the physics and detector are covered Chapter 4. The event selection and measurement is detailed in Chapter 5. Chapters 6 and 7 discuss the corrections and systematic uncertainties related to the energy flow measurement. The results are presented in Chapter 8, and Chapter 9 summarises and concludes the thesis.

2

Theoretical Overview

2.1 Quantum Chromodynamics

2.1.1 Fundamentals of Quarks and Gluons

The Standard Model (SM) describes the interactions of elementary particles observed till date. It accounts for the twelve known particles of matter (quarks and leptons) that interact via the strong, weak, and electromagnetic forces [33] [34] [35] [36]. The force-carriers, or particles responsible for the interactions are gauge bosons (gluons, W and Z bosons, and photons, respectively). Figure 2.1 (taken from [37]) lists the twelve fermions (red for quarks, orange and yellow for leptons) and four vector bosons (in blue) along with their spin, electric charge and masses. These particles combine to form the composite particles observed until now. The term particle could be considered a misnomer, however, given that the dynamics are governed by quantum mechanics. The particles exhibit wave-particle quality and are described by quantum state vectors in a Hilbert space. Consequently, the SM is a set of quantum field theories (QFT) [38] [39] that are used to characterise the behaviour of these fundamental constituents of matter.

The quantum field theory of the strong force, the fundamental force describing the interactions of quarks and gluons, is a non-Abelian, gauge QFT called Quantum Chromodynamics (QCD). According to the paradigm, quarks (spin-$\frac{1}{2}$ fermions) are the quanta of a Dirac-spinor field with a colour charge. Gluons are spin-1 bosons, the quanta of the colour field, born from the need for local gauge invariance of the quark fields. Together, the quarks and gluons (collectively referred to as partons) make up

hadrons such as the proton, neutron or pion. The property of colour in QCD is analogous to the electric charge in Quantum Electrodynamics (QED).

The colour field is characterised by the SU(3) symmetry of the QCD Lagrangian. Accordingly, there are three independent, conserved colour charges; red, blue, green. Like colour theory in the visual arts, mixing red, green and blue partons in equal proportions form a colour neutral object (white). Combining a parton with another that possesses the equivalent anti-colour produces the same result - formation of a "colourless" object. Quarks carry one unit of colour and gluons carry one unit of colour and anti-colour. Gluons carrying a colour charge is unique to QCD, arising from the theory's non-Abelian nature, and leads to self-coupling. However, they are electrically neutral, in contrast to quarks which carry $\frac{2}{3}$ or $-\frac{1}{3}$ e. A more in depth description of QCD can be found elsewhere [3] [36][40][41][42].

	1. Generation	2. Generation	3. Generation	Gauge Bosons
Quarks	u up $s = 1/2$ $Q = +2/3e$ $m = 1.5$–3 MeV	c charm $s = 1/2$ $Q = +2/3e$ $m = 1.27^{+0.07}_{-0.11}$ GeV	t top $s = 1/2$ $Q = +2/3e$ $m = 171.2 \pm 2.1$ GeV	γ photon $s = 1$ $Q = 0$ $m < 10^{-18}$ eV
	d down $s = 1/2$ $Q = -1/3e$ $m = 5.5$–6 MeV	s strange $s = 1/2$ $Q = -1/3e$ $m = 104^{+26}_{-34}$ MeV	b bottom $s = 1/2$ $Q = -1/3e$ $m = 4.2^{+0.17}_{-0.07}$ GeV	Z^0 $s = 1$ Q=0 $m = 91.2$ GeV
Leptons	e^- electron $s = 1/2$ $Q = -1e$ $m = 0.51$ MeV	μ^- muon $s = 1/2$ $Q = -1e$ $m = 105.6$ MeV	τ^- tau $s = 1/2$ $Q = -1e$ $m = 1777$ MeV	$W^\pm$ $s = 1$ $Q = \pm 1e$ $m = 80.4$ GeV
	ν_e electron neutrino $s = 1/2$ $Q = 0$ $m < 2$ eV	ν_μ muon neutrino $s = 1/2$ $Q = 0$ $m < 2$ eV	ν_τ tau neutrino $s = 1/2$ $Q = 0$ $m < 2$ eV	g gluon $s = 1$ $Q = 0$ m=0

Figure 2.1: Elementary particles of the Standard Model - The particles are listed according to the generation to which they belong, along with their corresponding spin, electric charge and mass. The Higgs-boson is not shown in this table. [37]

2.1.2 Key Aspects of QCD

Interactions described by QCD can be shown through simple Feymann diagrams, where any interaction between coloured objects can be drawn in terms of the fundamental

QCD vertices (see Figure 2.2). Diagrams which contain the fewest number of vertices for an interaction are termed Leading Order (LO) diagrams, while calculations at the next (higher) order are called Next-to-Leading-Order (NLO). Such diagrams may also contain closed loops, which have important consequences in all QFTs. Since the loops can be arbitrarily massive, diagrams containing them produce a divergence in calculations. These divergences are removed through the introduction of an artificial scale, μ_r, which limits the energy contained within loops. This method is called renormalisation.

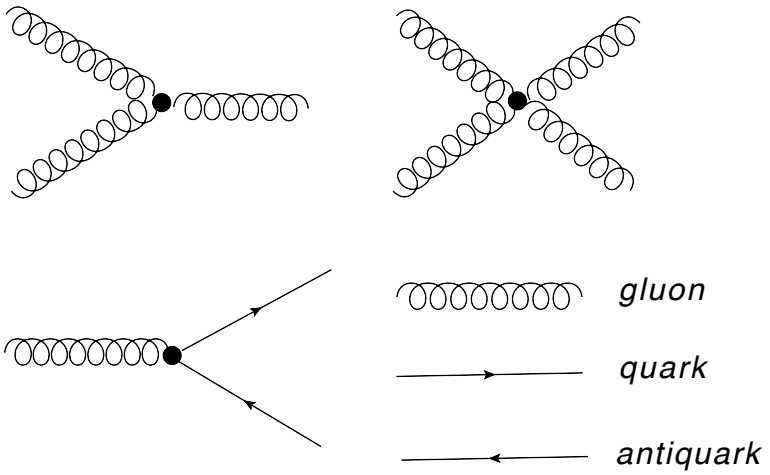

Figure 2.2: Fundamental vertices in QCD - $g + g \rightarrow g$ (top left), $g + g \rightarrow g + g$ (top right), $g \rightarrow q + \bar{q}$ (bottom left), and the legend (bottom right).

The dependence on μ_r and the divergences are absorbed into the coupling constant, α_s. As a result, α_s (and μ_r) is dependent on the scale of the process under study, earning it the name running coupling constant. At leading order, α_s is [36] [40] [42]:

$$\alpha_s(\mu^2) = \frac{4\pi}{\beta_0 ln(\mu^2/\Lambda^2_{QCD})} \tag{2.1}$$

where the Λ_{QCD} is a constant called the QCD-scale (~ 217 MeV) and μ is the energy scale. β_0 is a constant, the QFT beta-function which encodes the running of a coupling parameter, adapted for SU(3). Here, $\beta_0 = 11 - \frac{2}{3}n_f$, dependent on the number of flavours (n_f), where β_0 is positive at all energy scales. Therefore, the strength of the coupling decreases logarithmically with an increase in energy. This property is termed asymptotic freedom. The implication of asymptotic freedom is that at short enough distances, quarks and gluons de-couple and can be treated as quasi-free particles.

Equation 2.1 and Λ_{QCD} form the basis for two additional effects in observational QCD: confinement [43] and hadronisation [36]. When two colour charges separate, it becomes more energetically favourable for a new quark–antiquark pair to to be created from the vacuum, due to the large energy potential between the charges. The resulting colour-neutral particles cluster to form colourless hadrons (e.g. mesons and baryons). Confinement (or colour confinement) is the phenomenon whereby colour-charged particles cannot be directly observed and the clustering process is called hadronisation, or fragmentation.

Finally, the collinear properties of QCD such as collinear singularities and colour coherence must be considered. Generally speaking, they are the result of divergences in α_s at low energy scales. Colour coherence is the effect whereby hard gluon radiation tends to be emitted within a cone bounded by a pair of hard partons close to each other (colour dipole). At low energies, colour dipoles are more abundant because gluon emission is related to the scale of α_s. The profusion of gluons with low transverse energy causes the formation of a dense cloud of collinear gluons around a coloured object, known as a collinear singularity. Such phenomena have strong implications for both observational particle physics (e.g. jets 4.2.4) and in QCD calculations (e.g. factorisation of divergences 2.2.2).

2.1.3 Perturbative QCD

As in classical mechanics, the equations of motion for QCD are derived from its corresponding Lagrangian. Solving these equations analytically and making testable predictions is impractical, due to the infinite number of interactions possible. Numerical methods resolve this issue by using a perturbative expansion. At high energies, the coupling is small, allowing for a more manageable number of terms. Successive terms or orders of the expansion are proportional to higher powers of α_s. However, the validity of the expansion is limited; the scale of the interaction must be sufficiently large for the perturbative approach to be applicable. Given the impact of collinear effects and that the hadron structure is non-perturbative in nature, perturbative QCD (pQCD) is replaced by phenomenological models at the Λ_{QCD}-limit.

2.2 Proton-Proton Collisions

Proton-proton collisions are complicated processes due to the internal hadronic structure. The hadronic final state of hadron-hadron interactions can be described as the superposition of several contributions: products of the partonic hard scattering with the highest p_T, including initial and final state radiation; hadrons produced in additional multiple parton interactions (MPI); and beam-beam remnants (BBR) resulting from the hadronisation of the partonic constituents that did not participate in scatterings [44]. Figure 2.3 illustrates interactions at these three separate scales [45]. As such, interactions possess both hard and soft processes.

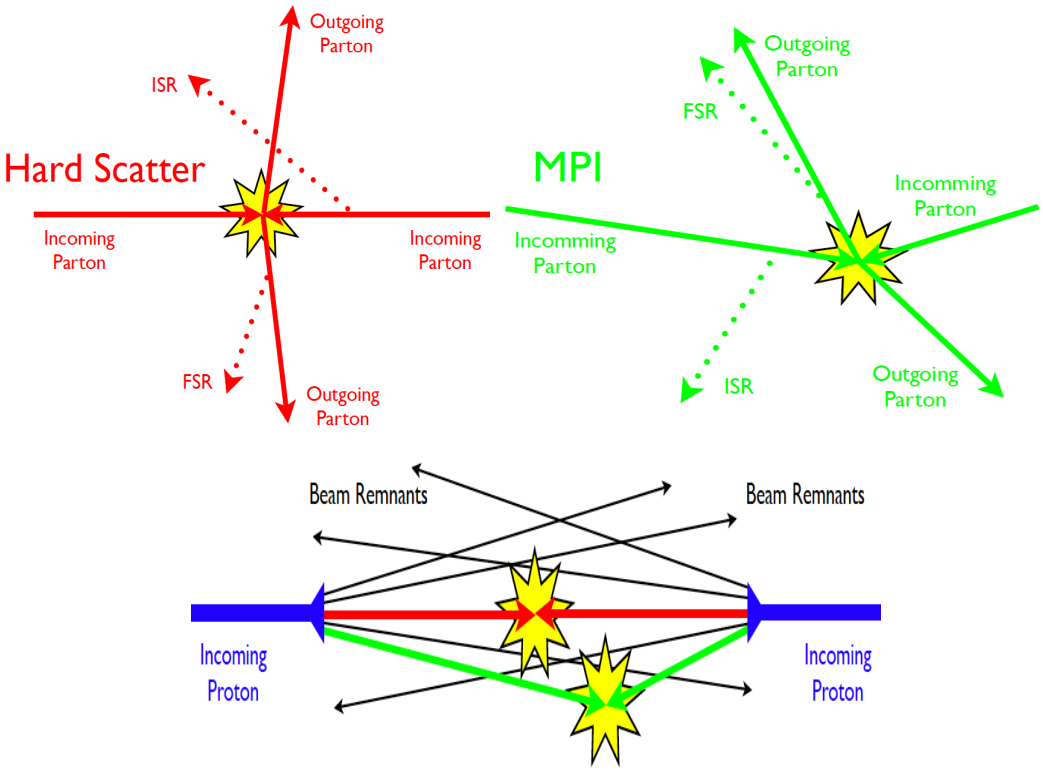

Figure 2.3: Components of a hadron-hadron collision - The final state of a hadron-hadron collision is a superposition of several contributions: products of the partonic hard scattering with the highest p_T (in red), hadrons produced in additional parton interactions (in green) and beam-beam remnants (in black) [45].

The hard (primary) interaction can be estimated through LO - and NLO-calculations

of the cross-section. However, the soft component cannot be treated in perturbation theory, as discussed previously. Moreover, the distribution of quarks and gluons within the proton has to be considered alongside the hard process [46]. Therefore, both pQCD and non-perturbative calculations are needed. The factorisation theorem separates the two, using pQCD to calculate the short-distance, process-dependent parton cross section, and determining the long-distance functions empirically. The long-distance functions include hadronisation, parton distribution functions fitted to data and models of multi-parton interactions. This section covers each of these aspects.

2.2.1 Structure of the Proton

Hadrons such as the proton contain two kinds of quarks, valence quarks (q_v) and sea quarks (q_s). Valence quarks contribute to the quantum numbers of the hadron and determine its behaviour, e.g., if it is a proton (uud) or a neutron (udd). Sea quarks are virtual quark-antiquark pairs which form upon the splitting of a gluon within the hadron. Conversely, the annihilation of sea-quarks forms gluons, resulting in a constant flux of q_s–g. The distribution of quarks and gluons within the proton is given by parton density functions (PDF)s, which take the form $f_{a|p}(x, Q^2)$. They give the probability to find a parton (quark or gluon) a, with the momentum fraction x, at the momentum scale Q^2 inside the proton. A more detailed discussion of PDFs is in Section 2.2.4.

2.2.2 Hard-Scattering Formalism: Factorisation Theorem

The factorisation theorem formulated by Drell and Yan [47] characterises the hard interaction between two protons. It postulates that the cross section for a particular process (σ_{AB}) can be reformulated as a convolution of the partonic cross section ($\hat{\sigma}_{ab}$) with the PDF of the proton ($f_{a|A}(x, \mu_F)$). The predicted cross-section is scale dependent, with the scale usually set to the same order of magnitude as the momentum scale of the hard interaction, i.e., α_s of the interaction.

As mentioned earlier, collinear singularities result in the perturbative calculation breaking down due to the divergences. As long as $\alpha_s < 1$, the perturbative series converges. This implies that pQCD describes gluon radiation as long as the scale is comparable to the momentum scale of the radiated gluon, but not low-p_T gluons emitted. In order to deal with the divergence, factorisation removes these singularities through the introduction of an arbitrary scale, μ_F, called the factorisation scale. The

scale, similar to the renormalisation scale (μ_R), creates an artificial limit between the divergent and convergent parts of the calculation. Taking into consideration the factorisation scale (μ_F) and the renormalisation scale (μ_R) of the hadronic cross section, the total cross-section is:

$$\sigma_{AB} = \int dx_a dx_b f_{a|A}(x_a, \mu_F^2) f_{b|B}(x_b, \mu_F) \left\{ \hat{\sigma}_{LO} + \alpha_s(\mu_R^2)\hat{\sigma}_{NLO} + ... \right\}_{ab \to X} \tag{2.2}$$

2.2.3 Partonic Cross Section

A matrix element is the amplitude of an interaction. The partonic cross-section, $\hat{\sigma}$, for a given hard interaction is calculated by integrating the square of the matrix element for a particular process over the amount of phase space available to the interaction. More often than not, the calculation is done numerically and divergences are avoided by excluding the corresponding regions of phase space. Leading order calculations give a good estimate of the expected cross-section. In the case of strong scale dependencies of the interaction, an NLO-calculation becomes necessary. They take into account virtual corrections, ISR and FSR. K-factor, the ratio between the LO and NLO cross section, quantifies the strength of the NLO calculation with respect to the LO.

2.2.4 Parton Distribution Functions

To obtain production cross sections at hadron colliders, the non-perturbative dynamics and pQCD calculations are merged by convoluting the matrix elements with the PDFs. PDFs give the distribution of quarks and gluons within the proton, parametrising the momentum fraction (x) of a parton within a proton. Through the factorisation theorem, collinear singularities are removed from the calculation of the partonic cross-section and absorbed into the PDFs. The factorised proton PDF is then applicable to all scatterings (ep, $p\bar{p}$ and pp) that adhere to the same factorisation convention.

Given that PDFs also contain the non-perturbative information about the hadronic interaction, pQCD cannot be used to derive them. Instead, they (starting distribution of a PDF) are extracted from data through global fits independent of the experimental set-up. With the starting distribution parametrised, the Q^2-evolution is treated with perturbative QCD via the DGLAP equations. Various groups such as CTEQ [48],

MSTW [49], and groups at HERA [50] have presented sets of PDFs based on data obtained from deep-inelastic scattering and related hard-scattering data from HERA and the TEVATRON. Different measurements are compared by evolving the parton distributions using evolution equations (see next section).

Figure 2.4 shows the MSTW PDFs for up-, down-, charm- and strange-quarks as well as for their anti-particles. The gluon distribution, scaled down by a factor of 20, shows a significant rise for low-x, indicating an increase in gluon density within the proton, while the valence quarks are dominant at high-x values.

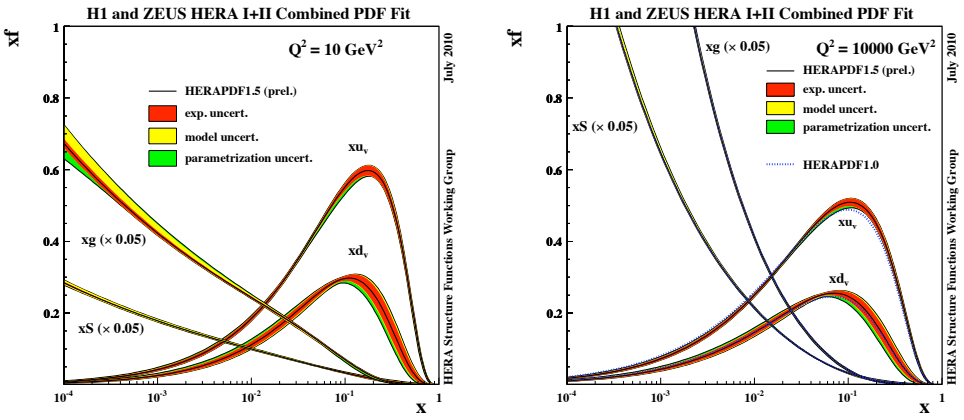

Figure 2.4: HERA **parton distribution function at** NLO - $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 [51]. The gluon distribution has been scaled down by a factor of 20.

2.2.5 Parton Evolution

Perturbative QCD cannot predict the internal distributions of partons within a proton, but it can model how the distributions evolve for a given input (e.g. at a given scale, Q^2). The x-dependence of PDFs are obtained from fitting experimental data, while the development of the partons can be described by one of the many parton evolution schemes that exist. The Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) equations [52], [53], [54], [55] determine the distributions for higher Q^2, starting from an initial distribution at a predefined scale. In this framework, the parton cascade is strongly ordered in transverse momentum (k_T), with only a weak kinematical-ordering for the momentum fraction (x).

However, at large centre-of-mass energies, $\log(1/x)$ contributions are non-negligible, limiting the application of DGLAP to the scenario where $log(\mu) \ll log(\frac{1}{x})$. The Balitsky-Faden-Kuraev-Lipatov (BFKL) evolution equations [56], [57], [58], [59] resum the $\log(1/x)$ contributions and is therefore believed to be the dominant scheme at low-x. The cascade follows a strong ordering in x but not in k_T, which implies that a clear distinction between DGLAP- and BFKL-dynamics should be observable at low-x regions. Yet, no conclusive evidence for BFKL-physics has been observed till date and therefore BFKL equations are not used for evolving extracted parton distributions (PDFs).

The DGLAP equations set the scale dependence and the PDF is written as:

$$\frac{df_{a|p}(x,Q^2)}{dlogQ^2} = \frac{\alpha_s}{2\pi}\int_x^1 \frac{dz}{z}\left\{P_{aa'}(z',\alpha_s)f_{a'|p}(\frac{x}{z'},Q^2)\right\} \tag{2.3}$$

where $P_{aa'}(z,\alpha_s)$ gives the probability that a parton a evolves into another parton a' carrying the momentum fraction z by emitting a gluon or quark. Known as the Altarelli-Parisi splitting functions [54], they describe the amplitudes of the processes that generate additional partons. They can be expanded in α_s as follows:

$$P_{aa'}(z,\alpha_s) = P^{LO}_{aa'}(x) + \frac{\alpha_s}{2\pi}P^{NLO}_{aa'}(z) + \left(\frac{\alpha_s}{2\pi}\right)^2 P^{NNLO}_{aa'}(z) + ... \tag{2.4}$$

Figure 2.5 shows the elements that form the basis of initial-state (ISR) and final-state (FSR) radiation, described by equation 2.4. ISR showers are QCD emissions (radiation) from the initial-state partons and radiation from final-state partons is known as FSR. This means that the internal structure and evolution of the incoming partons is related to the ISR while the fragmentation of the outgoing partons is dependent on the FSR. Experimentally, radiation emitted before the hard interaction and close to the beam line is considered as ISR and FSR occurs after the hard interaction in close proximity to the outgoing partons.

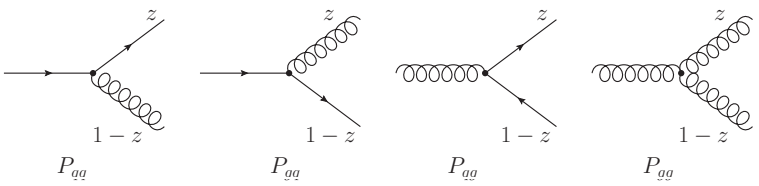

Figure 2.5: Altarelli-Parisi splitting functions - The leading order Altarelli-Parisi splitting function diagrams [60].

The Ciafaloni-Cantani-Fiorani-Marchesini (CCFM) scheme [61] [62] attempts to describe both the BFKL (low) and DGALP (high) dominated x-regions. A LO evolution equation, CCFM is based on the principle of colour coherence, i.e., emissions of gauge bosons. As a result, it features only gluon emissions instead of quarks (e.g. gluon jets). A description of the scheme and its hadronic final-state predictions can be found in [63].

2.2.6 Underlying Event in Proton-Proton collisions

In hadron-hadron collisions, the final states of a hard interaction can be described as the superposition of several contributions: products of the partonic hard scattering plus initial- and final-state radiation, hadrons produced in additional multiple parton interactions (MPI) [64], and beam-beam remnants (BBR) resulting from the hadronisation of the partonic constituents that did not participate in any scattering. The combination of MPI, BBR, ISR and FSR is called the underlying event.

A primary process is defined as a parton-parton interaction which is completely insensitive to additional incoming particles and beam remnants, including coherent radiation of all orders; the source of the partons is irrelevant, i.e., the definition holds true for ep, pp and $p\bar{p}$ collisions. From this stems a working definition for the underlying event (UE): all the activity or energy flow that is not associated with the primary process. However, it excludes contributions from pile-up events as they are beam-luminosity dependent and not characterised by the physics of an event.

A predominant source of underlying event activity is multi-parton interactions. When composite particles collide, more than one pair of partons can interact. Given the rise of parton densities at low x, the probability for additional parton-parton scatters increases at such regions (illustration in Figure 2.6). These secondary collisions are called multi-parton interactions. They can occur at both ep and hadron-hadron colliders, contributing significantly to an increase in the interaction cross section. Other sources of UE activity include initial and final state radiation, multiple scattering where a primary particle re-scatters off the remnants, beam-beam remnants, and other effects that may affect primary scattered partons.

Fragmentation models deal with the hadronisation of the final state particles and parton shower models describe gluon radiation in the initial and final state. Implemented in event generators, these models have been tuned to describe data taken at

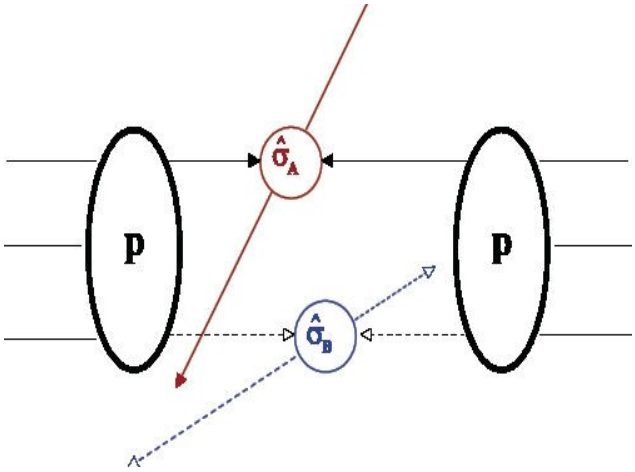

Figure 2.6: Increase in cross-sections due to MPI - With an increase in the density of partons with decreasing x, the probability for more than one partonic interaction per event increases. Multi-parton interactions lead to increased cross-sections.

previous colliders such as LEP, HERA and the TEVATRON. Similarly, multi-parton interaction models, have made a lot of progress recently. There is a large amount of experimental evidence for MPIs in hadron-hadron and ep interactions; measurements at ZEUS [60], CDF [45] [65] [66] [67] and H1 [68] [69][70], have provided some insight into their nature.

MPIs may be very soft (i.e. small transverse momenta) constituting the soft underlying event, or hard, i.e., with high enough transverse momenta to form jets. Signals of such activity include an increase in charged particle multiplicity and an E_T pedestal, a surge in the production of mini-jets, and an augmented jet rate and energy scale. Moreover, events with high p_T jets in the final state show a higher UE activity. The discrepancy between the data and event simulations without MPIs is accentuated for multi-jet events (e.g. 4-jet events, necessary for top mass reconstruction) [71]. In fact, there is a need for more than a fifty percent double-parton interaction to describe data in certain experiments [72]. Finally, MPIs impact triggering strategies as the additional activity can influence isolation criteria (e.g. Drell-Yan measurements) and reduce the rapidity gap survival probability (e.g. diffraction). As a result, they contribute significantly to jet (and possibly Higgs) production and can adversely affect the analyses at collider experiments.

To understand this QCD background for the primary interactions, a coherent description for high and low p_T multi-parton interactions is needed [73]. With the LHC operating at far higher centre-of-mass energies than HERA and the TEVATRON, the need to understand such phenomena is imperative for future measurements the TeV scale. Phenomenological models attempt to describe MPIs, using different parametrisations of the correlations between momentum and energy conversation, and other quantum effects. However, the extrapolations of these predictions to LHC energies is unreliable, motivating the need for further investigations.

Figure 2.7 depicts the complexity of a hadron-hadron event with a primary hard scatter ($W^+ \rightarrow c + \bar{s}$) and a secondary or semi-hard interaction (recoiled d-quark) at the top of the picture, gluon radiation (ISR and FSR) in the middle, and soft MPI ($gg \rightarrow gg$) overlaid on it at the bottom of the picture.

2.3 Physics in Event Generators

Event generation and simulations are key to most analyses in particle physics. The different components of event generation used in this analysis and the way they come together is covered in Chapter 4. In this section, the implementation of QCD-physics in the event generators is briefly discussed. The key differences in hadronisation procedures, modelling of parton showers and MPIs are highlighted via the comparison of two multi-purpose generators, HERWIG [75] and PYTHIA [76]. Where relevant, alternative methods employed by programs such as CASCADE [63] and SYBILL are expanded upon. The concepts introduced previously and discussed here are central to the data-simulation comparisons in Chapter 8.

2.3.1 Minimum p_T

The integrated cross section can be written as a function of a low transverse momentum scale ($p_{T,min}$), such that:

$$\sigma_{int}(p_{T,min}) \propto \frac{1}{p^2_{T,min}}, \tag{2.5}$$

If the $p_{T,min}$ scale is infinitely small, the integrated cross section above the $p_{T,min}$-threshold diverges and exceeds the total pp cross section [77] [78]. To suppress scatters

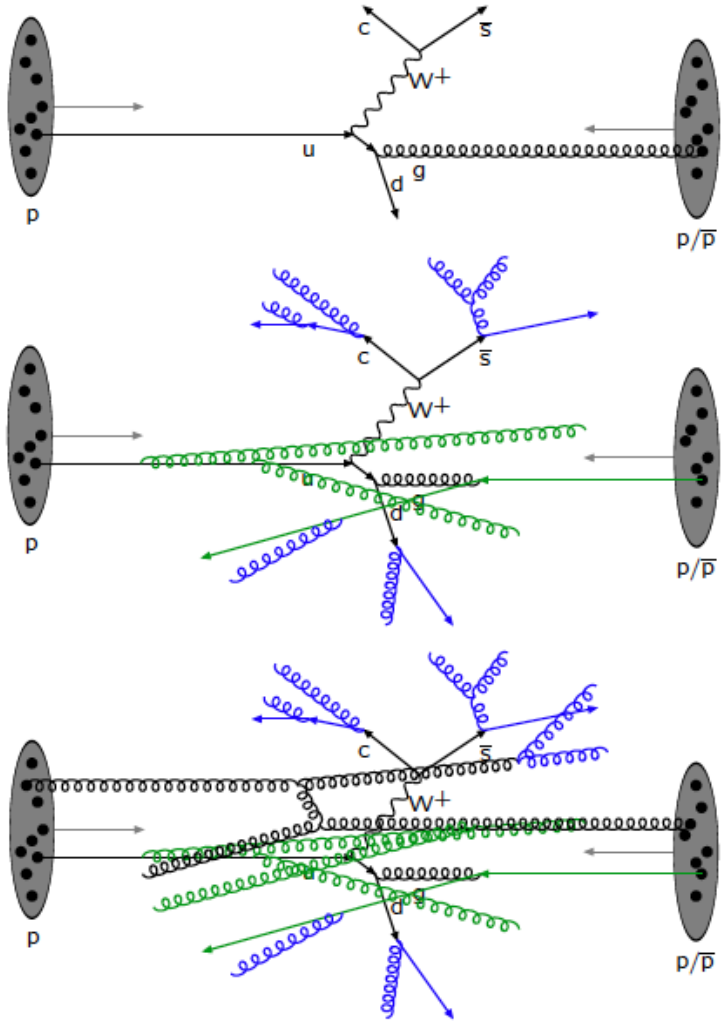

Figure 2.7: Pictographic representation of proton-proton collisions at high energies - Hard process plus a secondary scatter (top), overlaid with initial- and final-state radiation (middle), and finally a soft MPI involving radiated gluons added to the event (bottom). [74]

below $p_{T,min}$, a lower limit on the minimum transverse momentum of the interaction is imposed. Since the inclusive cross section scales as $1/p_T^4$ and high-p_T states are rare, discarding events below this cut-off also improves the efficiency of generating high p_T events. Nonetheless, to ensure an unbiased event sample with a realistic amount of soft activity, $p_{T,min}$ is kept moderately small. HERWIG employs an abrupt $p_{T,min}$-cut-off whereas PYTHIA utilises a smoothly dampened, adjustable variable. More on the variation of this parameter and its impact on MPIs can be found in Section 2.3.4.

2.3.2 Parton Showers

Parton showers play a key role in event simulations by approximating the amount of ISR and FSR in a collision. While the showers for ISR and FSR are simulated differently because their cascades are space-like (ISR) and time-like (FSR), the general approach remains the same. DGLAP equations relate partons from the hard interaction to partons at the Λ_{QCD} scale, describing the development, kinematics and flavours of the emitted partons. The development is characterised by an evolution variable; in PYTHIA, it is either the virtuality of the parent parton (Q^2) or the transverse momentum squared (k_T^2) of the two emitted partons. In HERWIG, the evolution variable ($E^2(1-\cos\theta)$) depends on the energy (E) of the parent parton and the emission angle (θ) between the two showered partons. The angular-ordered parton showers in HERWIG account for the ordering present in cascades due to colour coherence; an additional, adjustable constraint is introduced in the PYTHIA shower model to describe its impact. Infrared singularities cancel out in both approaches.

2.3.3 Hadronisation and Fragmentation

Once the showering cascade is finished, i.e. the virtuality of the partons $Q_0^2 < 1GeV^2$, hadronisation models cluster the emitted partons. Two iterative models are predominantly used, the cluster fragmentation model (HERWIG) [79] [80] and the Lund string model [81] [82].

The Lund string model assumes colour flux tubes or strings between $q\bar{q}$ pairs exiting the parton shower. The tubes are defined such that they are uniform in length, have negligible transverse momentum and posses the diameter of a meson (1 fm). The behaviour of the tubes is based on the principles of confinement. As the $q\bar{q}$ pairs move apart, the colour strings are stretched until it is more energetically favourable

to fragment and create an additional $q\bar{q}$ pair. The procedure is repeated until all the strings are too light for further fragmentation.

The cluster fragmentation model follows a simpler strategy to combine partons into observable hadrons. Final-state gluons are split into $q\bar{q}$ pairs which are grouped with the other outgoing quarks and diquarks to form colour-singlet clusters. Light quark-antiquark pairs are added to the cluster, fragmenting them until stable hadrons are formed. In cases where a cluster is too light to fragment into a hadron, a light hadron replaces the cluster.

2.3.4 Multi-Parton Interactions

The cross section for the hard (primary) interaction, given by Eq. 2.6, diverges faster than $1/p^2_{T,min}$ for $p^2_{T,min} \to 0$. At low values of $p_{T,min}$, σ_{hard} eventually exceeds the total inelastic cross section, as seen in Figure 2.8. This difference between the two calculations can be explained if σ_{hard} is interpreted as the partonic cross section instead of the hadron-hadron cross section; it contributes multiple times to the integrated cross section. However, σ_{hard} contributes only once to the total non-diffractive cross section (σ_{nd}). Therefore, the average number of parton-parton scatterings above $p_{T,min}$ in an event can be given by Eq. 2.7.

$$\sigma_{hard}(p^2_{T,min}) = \int_{p^2_{T,min}}^{s/4} \frac{d\sigma_{hard}}{dp^2_T}\, dp^2_T\,, \tag{2.6}$$

$$\langle n \rangle = \frac{\sigma_{hard}(p_{T,min})}{\sigma_{nd}}, \tag{2.7}$$

Initial measurements of charged particle multiplicity at UA5 [83] demonstrated that data is better described by accounting for additional activity in an event. From this emerged a simple model for multi-parton interactions [64]. Subsequent measurements of charged particle multiplicity and γ+three-jet events at CDF [84], and multi-jet events in photoproduction at HERA provided more insight into this phenomenon by probing specific regions especially sensitive to MPI activity.

Though the data paved the way for old models to be adapted and new models to be formulated, the fundamental concept behind modelling MPIs remained the same. That is, to apply the QCD factorisation theorem to the additional scatters as well, with the transverse momentum of the additional scatter ($p_{T,as}$) adopted as the factorisation

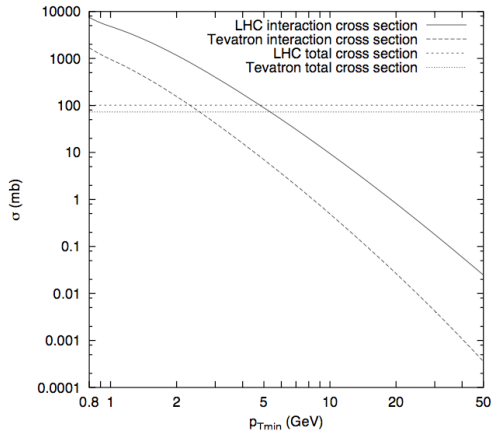

Figure 2.8: Integrated interaction cross section - The integrated interaction cross section (σ_{int}) above $p_{T,min}$ for the TEVATRON with 1.8 TeV $p\bar{p}$ collisions, and the LHC, with 14 TeV pp ones. The flat lines represent the respective total cross section. Taken from [78]

scale ($\mu_F = p^2_{T,as}$). Therefore, the basic form of the second interaction's cross section is a partonic cross-section convoluted with a parton density.

The number of interactions is strongly dependent on the minimum transverse momentum cut-off ($p_{T,min}$). With a lower value corresponding to more multiple interactions, there is also an implicit dependence between the cut-off and the average charged particle multiplicity ($dN_{chg}/d\eta$) in an event. Thus, at any energy scale, the simplest way to determine $p_{T,min}$ is to impose a requirement that the predicted particle multiplicity agree with data.

Extracting $p_{T,min}$ from data inexorably leads to the question of the variable's energy-dependence. Relevant for comparisons of jet rates at different collider energies, the issue is in the limelight with the start-up of the LHC. Post HERA data, PDFs predict parton densities to be rising more steeply at small-x (e.g. x-regions explored at the LHC) than initially assumed during the conception of the simple MPI model. The default energy dependence given by:

$$p_{T,min}(s) = (1.9GeV)\left(\frac{s}{1\ \text{TeV}^2}\right)^{\epsilon}, \tag{2.8}$$

such that $p_{T,min}$ is assumed to increase in the same way as the total cross section; that is, with some power (here, $\epsilon \approx 0.08$) which relates to the behaviour of parton distributions at small-x and Q^2 [76]. Since the $p_{T,min}$ values are determined for given

model parameters and a given calculation of jet cross sections, the parameter must be retuned according to each scenario.

Collisions are also parametrised by impact parameters, b, which quantify the overlap of matter in hadrons. Most models use such impact parameters but vary in the assumptions made on the matter distribution within the hadrons. The greater the overlap, the more likely it is for additional interactions to occur. Therefore, MPI models include impact parameters that can be adjusted. Models which consider an impact-parameter dependence of MPI also use the same ansatz given in Eq. 2.12 for the $p_{T,min}$ cut-off's energy-dependence, but with additional tweaking (e.g. 1.9 GeV $\rightarrow$ 2 GeV).

Other aspects of a collision that may impact MPI [64] and vice versa include parton showering, possible scenarios for colour string connections, saturation effects, and diffraction in MPI events.

The colour flow in an event can affect the angular distribution of gluon radiation. Moreover, the colour structure can change the behaviour of soft, background partons. In the MPI models developed initially, simple assumptions were made on colour flow, with colour connections of the beam remnant only to the first hard scattering. Similarly, early models included a treatment of initial-state and final-state parton showers, but only for the hard scattering. Subsequent improvements in the treatment of parton showers, specific to the two most commonly used event generators, is discussed in the following sub-sections.

The connection between MPI, saturation and diffraction is given via the Abramowski-Gribov-Kancheli (AGK) cutting rules [85]. Though the treatment forms a basis for the PHOJET event generator [86] (used briefly in Chapter 7), its impact is not considered an integral part of this study. As a result, the relation will not be discussed here. More on the subject of AGK-rules and their role in MPIs can be found in the following references [87] [88] [89].

It is important to remember that MPIs are the least understood aspect of hadron-hadron collisions. Despite the immense progress made in both phenomenological descriptions and experimental techniques over the last twenty years, further measurements to investigate MPI-behaviour are necessary. As such, there are few observables and event topologies which are sensitive to the physics of multi-parton interactions; charged particle multiplicity in minimum bias events, jet pedestals in the transverse region for high p_T jet events, three-jet+γ and four-jet angular de-correlations have

proved useful in studies of the semi-hard component of the underlying event. Models have been developed as a consequence of such analyses, and tuned to data (e.g. CDF [90]). The following sections discuss the latest features of MPI models in the two most widely-used event generators, PYTHIA and HERWIG.

2.3.4.1 PYTHIA

PYTHIA utilises a smoothly dampened, adjustable $p_{T,min}$ to place a lower limit on the minimum transverse momentum of the interaction, where $p^2_{T,min} \rightarrow p^2_{T0} + p^2_T$ and p_{T0} is modified to describe data. The additional interactions are ordered (high to low) by their transverse momenta, before calculating the probability of their occurrence. For each scatter, activity at lower p_T scales are integrated out of the parton densities; the single-parton densities are re-scaled to account for flavour and momentum correlations between partons. The generator also uses an impact parameter, b, tuned to data. The impact parameter is based on different choices for the overlap function (e.g. single Gauss, exponential, double Gauss) [78]; by default, the matter is distributed according to a spherically-symmetric double-Gaussian. The formalism for MPI assumes the probability for n interactions for a given impact parameter b is characterised by a Poisson distribution, as it considers each parton scattering to be independent from each other.

Finally, the latest MPI model features interactions interleaved with parton showers [91], i.e., it treats initial-state parton showering and MPI at the same time. New, transverse-momentum-ordered showers are introduced, making p_T the common evolution scale for multi-parton interactions, and for initial-state and final-state radiation [76]; Eq. 2.9 formulates the probability distribution for an interaction at a given p_T scale, composed from its various contributions. Each multiple interaction is associated with its set of ISR, as shown in Figure 2.9.

$$\frac{dP}{dp_T} = \left(\frac{dP_{MPI}}{dp_T} + \sum \frac{dP_{PS}}{dp_T} \right) \otimes \exp \left(\frac{dP_{MPI}}{dp_T} + \sum \frac{dP_{PS}}{dp_T} \right) \tag{2.9}$$

The figure illustrates the structure of a hadron and the evolution of its partons; it depicts one hard interaction and three subsequent interactions at successively lower p_T scales, along with their associated initial-state radiation. The approach is to first consider the hardest interaction ($p_{\perp,1}$), thus allowing the use of normal parton densities for the initial scattering. Next is the description of the second interaction ($p_{\perp,2}$),

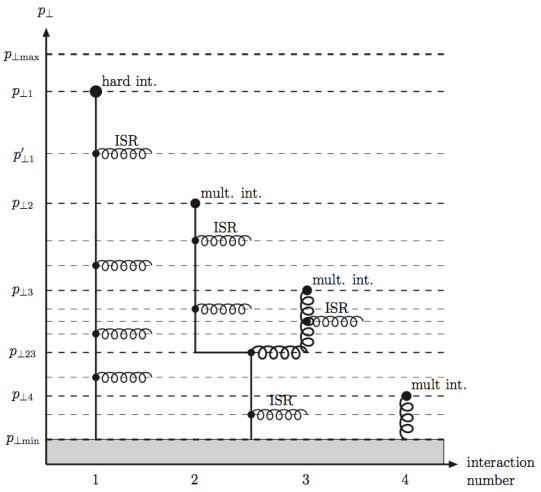

Figure 2.9: Schematic Representation of Interleaved Parton Showers and Multi-parton Interactions - Schematic figure illustrating one hard interaction and three subsequent interactions at successively lower p_T scales, with associated initial-state radiation. Full lines and spirals represent quarks and gluons, respectively. Taken from [91]

followed by the soft ISR associated with the hardest interaction. Giving priority to the second interaction over ISR associated to the hardest interaction is what allows the use of the standard parton densities to characterise the primary interaction ($p_{\perp 1}$) and the ISR that occurs above $p_{\perp 2}$. The known (or estimated) correlation effects are introduced via the subsequent scatterings at lower p_T. The procedure is repeated and the evolution continues until $p_{\perp,min}$, the non-perturbative border or the cut-off point is reached. A complete picture would include a mirrored image for the other hadron, with the two incoming hadrons joined at the interactions.

2.3.4.2 HERWIG

Unlike PYTHIA, HERWIG imposes a sharp cut-off for $p_{T,min}$, the additional scatters are generated with matrix elements similar to the primary interaction, and the single-parton densities are not re-scaled. Moreover, the MPIs are not ordered in p_T nor interleaved with parton showers. Finally, instead of PYTHIA's impact-parameter functions (e.g. double-Gaussian), HERWIG adopts the eikonal paradigm [92] to characterise the overlap between protons. The use of the eikonal model is based on the assumption that for a fixed impact parameter (b), the individual scatterings are independent and that distribution of partons can be characterised by b and their x-dependence [92] [93].

The model approximates interacting hadrons as disks extended in the transverse plane [92]. The impact parameter, b, quantifies the overlap of the two disks during a collision. A function $A(b)$ (cf. Eq 2.10) dependent on the impact parameter provides the distribution of the partons within the interacting hadron; it is modelled by the electromagnetic form factor, the inverse proton radius (μ), and a modified Bessel function (K_3). Together with the inclusive cross section σ_{inc}, the partonic overlap function of the colliding hadrons ($A(b)$) gives the average number of partonic collisions for a given value of b via Eq 2.11. Here, a large impact parameter and dense distribution implies a larger probability for MPIs. The number of scatters, n, are distributed according to a Poissonian, due to the assumed uncorrelation between them.

$$A(b;\mu) = \frac{\mu^2}{96\pi}(\mu b)^3 K_3(\mu b)\,, \tag{2.10}$$

$$\langle n(b,s)\rangle = A(b)\,\sigma_{inc}^{hard}(s;p_{T,min}) \tag{2.11}$$

Though the assumption that the scatters are independent is a valid approximation for small-x regions, there is a fraction of events for higher x values that lead to some correlations. By vetoing any interaction which would violate the conservation of total momentum, flavour and energy of the event, the correlations are taken into account.

The generation of additional scatters occurs only after the hard scatter and its associated parton showers have been simulated. The number of multi-parton interactions is determined via the formalism described earlier, and are generated in a manner similar to the primary interaction. That is, a matrix element calculates the additional scatter and parton showers evolve the produced particles for each of these interactions. Upon evolution, if the process violates conservation principles, the scatter is vetoed. Centre-of-mass dependent parametrisations and colour flow reconnections for MPI are a relatively new feature; details can be found in [94].

2.3.5 Brief Description of Monte Carlo Models and their Tunes

Parameters in event generators can be adjusted such that the predictions of physics at non-pertubative scales agree with experimental data. These parameter sets differ in their choice of flavour, fragmentation and UE properties. Underlying event tunes, i.e.,

settings relevant to this measurement include specifications on the minimum p_T-cut-off and $\sqrt{s}$-energy dependence, impact parameters, possible momentum correlations, parton showering and choice of PDFs. In this section, the main features of the Monte Carlo models used for comparison with the data (see Chapter 8) are presented.

2.3.5.1 PYTHIA6.4

The primary generator used in this analysis, for both the comparisons and the correction of the data is PYTHIA6.420 [76]. An average of predictions by the tunes D6T [66], ProQ20 [95], Z2, and P0 forms the basis of the final correction factor (see Chapter 6). Additional parameter sets compared to corrected data include DW [66], Pro-pT0 [95], Perugia 2011 (P11) [96] and the ATLAS Minimum Bias Tune 1 (AMBT1) [97].

The tunes differ in how they regularise the formal $1/p_T^4$ divergence of the leading order partonic scattering amplitude at the limit $\hat{p}_T \to 0$. As discussed earlier, the divergence is regularised by the introduction of a cut-off parameter (p_{T0}), which is energy-dependent; the parameter's energy dependence shown in Eq. 2.12 can be rewritten as:

$$p_{T0}(\sqrt{s}) = p_{T0}(\sqrt{s_0}) \left(\frac{\sqrt{s}}{\sqrt{s}_0} \right)^{\epsilon}, \tag{2.12}$$

where $\sqrt{s}_0$ is the reference energy at which p_{T0} is determined, and ϵ some power which relates to the behaviour of parton distributions at small-x and Q^2. Since the value of p_{T0} governs the amount of multi-parton interactions, the different values will produce varying amounts of additional activity. For example, D6T describes the lower energy UA5 and TEVATRON data, but is known to predict lower UE activity with respect to the other tunes [98]. The other PYTHIA6 tunes which adopt the TEVATRON reference energy (e.g. ProQ20 and P0) also fail to describe $\frac{dN_chrg}{d\eta}$ and UE data at the LHC. The UE settings for P11 and AMBT1 are based on the Professor tunes [95] and recent measurements at the LHC, including charged multiplicity [99] and UE [44] [98] data.

The tune Z2 is similar to Z1 [100], featuring p_T ordered parton showers. The difference between the two sets is that the transverse momentum cut-off at the nominal energy of $\sqrt{s_0} = 1.8$ TeV is decreased by 0.1 GeV/c for Z2. Moreover they differ in their choice of PDF. While CTEQ5L [101] is the default for the majority of the PYTHIA6 tunes, D6T and Z2 prefer to implement the newer PDF set CTEQ6L [102] to describe the

protons. Both CTEQ5L and CTEQ6L are PDFs obtained in a global fit with LO hard cross sections. However, the CTEQ6L is a newer generation of parton distribution functions with increased precision and quantitative estimates of uncertainties. Moreover, the new gluon distribution is considerably harder than that of previous standard fits [102]. The majority of them also share common flavour and fragmentation parameters, determined using data from LEP. Two models which feature improved high-z fragmentation, where z denotes the ratio of the momentum of an outgoing charged particle to the momentum of the parton initiating the hard scatter, are P0 and ProQ20. Finally, P11 and Z2 feature a new MPI model [103] interleaved with the parton showers. A more detailed comparison of the tunes can be found in [96].

2.3.5.2 Alternative generators

PYTHIA8 [104] also incorporates the new MPI model but the predictions correspond to the default settings of the generator. A unique characteristic of this version is that it introduces a simulation of the hard diffractive component in an event, distinguishing it from most other generators including HERWIG. The HERWIG++2.5 [93], used for comparisons with data, has been tuned independently for different centre-of-mass energies. Parameters for the UE and colour reconnections [94] were tuned to UE and minimum-bias data at $\sqrt{s}$=0.9 TeV and to UE data at $\sqrt{s}$ =7 TeV [105]. It should be noted that the precise description of diffraction is of little relevance for this study, since the diffractive contributions are strongly suppressed by the event selection requirements (see Chapter 5 and Chapter 6). Therefore, it is reasonable to expect a better description of the corrected data by the tuned HERWIG++2.5 models, than the default PYTHIA8 tune where predictions include diffraction but are not adjusted to describe LHC data.

In contrast to PYTHIA and HERWIG where parton showers are based on the DGLAP evolution equations, DIPSY is an event generator [106] based on a dipole picture of BFKL. Predictions from it are relevant for this measurement since it is built to describe parton dynamics at low-x and non-diffractive final states. Another generator with an alternative description of parton showers is CASCADE [63], which implements the CCFM evolution equations for the initial-state cascades, supplemented with off-shell matrix elements for the hard scattering. DIPSY includes a model for MPI while CASCADE does not; neither generator is tuned to describe recent data from hadron colliders.

2.3.5.3 Cosmic-ray event generators

Lastly, predictions from hadronic MC event generators based on Reggeon-Field-Theory (RFT) [107] are considered. In general, the soft component is described via exchanges between virtual quasi-particle states (pomerons), which are governed by RFT; the soft interactions include elastic, diffractive and soft-inelastic scatterings [108]. Semi-hard interactions (e.g. $\sqrt{s} \approx \mathcal{O}(100\text{ GeV})$) can be treated within pQCD, with partonic cross sections convoluted with PDFs; the perturbative parton splittings are described by DGLAP equations while the non-perturbative effects are modelled using the Lund string model and beam-remnant modelling [108] [109]. For energies comparable to those at the LHC, additional effects such as saturation effects and MPIs also play a role - the interaction cross section exceeds the total inelastic cross section and low-x effects (rise in gluon distributions) become increasingly important. Such effects are accounted for by employing eikonal (multi) parton ladders and a regime for treating gluon saturation [110]. The perturbative interactions use RFT generalised to treat parton-parton interactions via hard "cut" Pomerons ("semi-hard Pomeron" approach [111]), or the DGLAP formalism (pQCD) according to the so-called minijet model [112].

The approach described above is used primordially in event generators built for cosmic-ray physics. Predictions from EPOS 1.99 [113], QGSJET II [114], QGSJET 01 [115] and SIBYLL [116] are compared to corrected data in Chapter 8.

EPOS takes into account energy-momentum correlations between multiple re-scatterings, with the description of non-linear effects is based on an effective treatment of lowest order pomeron-pomeron interaction graphs. QGSJET-II also treats non-linear parton effects via pomeron interactions, but they are based on an all order re-summation of the corresponding Reggeon-Field-Theory diagrams. Based on the dual parton model (DPM) [117], SIBYLL also employs Pomeron formalism, but uses the minijet model for the perturbative interactions and the Lund model for hadronisation. In contrast to SIBYLL, semi-hard scattering processes are treated as exchanges of "semi-hard" Pomerons in QGSJET. The models diverge in their predictions primarily due to different treatments of non-linear interaction effects related to parton shadowing and saturation [108] [109] [110].

The model parameters are tuned with accelerator data from SppS and the TEVATRON. In general, 0.9 TeV data at the LHC is well-reproduced but the particle mul-

tiplicity is not well predicted at 7 TeV [108]; the simplest models (QGSJET, SIBYLL) seem to provide the best description.

2.4 Relevance of the Measurement

2.4.1 Predictions for the LHC

The procedure outlined in Section 2.2 has been used to calculate SM cross sections at NLO, for the TEVATRON and the LHC. The results are shown in Figure 2.10, where the two vertical lines denote the TEVATRON operating $\sqrt{s}$-energy for $p\bar{p}$ collisions and the LHC's design $\sqrt{s}$-energy for pp collisions.

However, additional interactions (i.e., MPIs) lead to an increase in measured cross-sections, leading to a significant increase in background for important signal processes. At large collision energies, the momentum fraction of the proton (x) carried by the partons can become very small and the parton densities increase, giving rise to multiple parton interactions. Thus, the probability of more than one partonic interaction per event increases for collision energies produced at the LHC, with respect to the TEVATRON.

Till date, measurements of the underlying event structure at hadron colliders have been performed for central pseudorapidities, i.e., $|\eta| < 3$, where $\eta = -\ln\left[\tan\left(\theta/2\right)\right]$, θ denoting the polar angle of the particles with respect to the beam axis [44, 65]. Since two units in pseudorapidity roughly translate to one order of magnitude in x [119], analysing data from the forward regions allows us to probe the partonic content of the proton at smaller x values. With this increase in phase space, we can further examine the influence of the evolution process on the hard interaction, the behaviour of the proton remnant and the fragmentation process. This will allow us to further constrain existing underlying event models. Furthermore, measurements could provide first indications of alternative parton dynamics such BFKL or CCFM evolutions. Figure 2.11 shows the accessible (x, Q^2) region for a final state with invariant mass M (e.g. mass of muon pair, Z-boson, or top quark), the factorisation scale equal to M, and rapidity y, at the LHC for two collision energies at the LHC, $\sqrt{s} = 7$ TeV (left) and $\sqrt{s} = 14$ TeV (right).

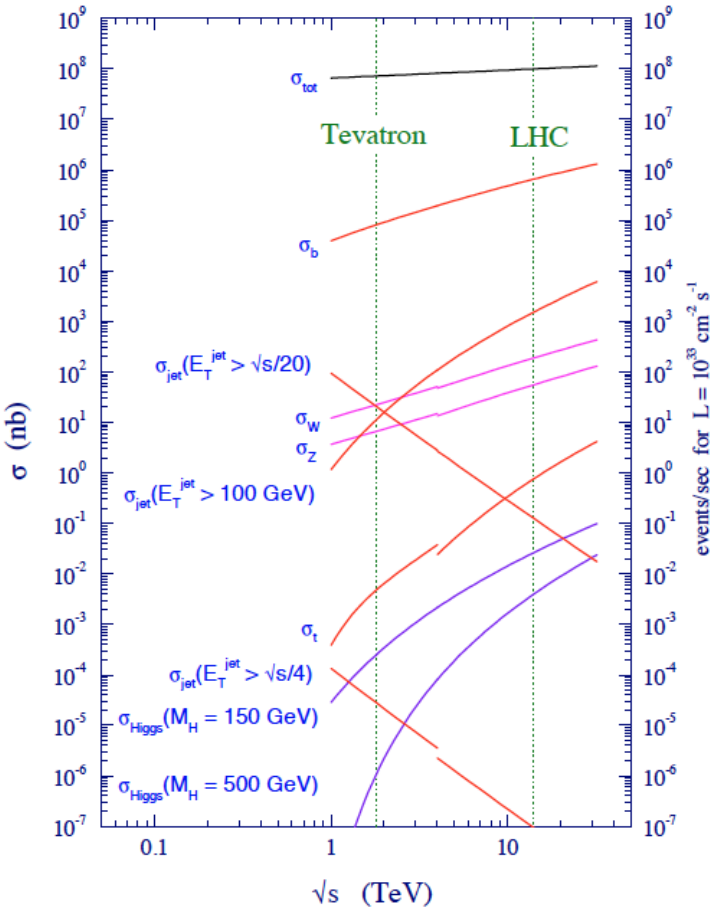

Figure 2.10: Standard Model cross sections at the TEVATRON and the LHC - Cross sections for various hard-scattering processes at the TEVATRON and the LHC, calculated at NLO, using the MRS98 PDF. The two vertical lines denote the TEVATRON operating $\sqrt{s}$-energy for $p\bar{p}$ collisions and the LHC's design $\sqrt{s}$-energy for pp collisions. [118]

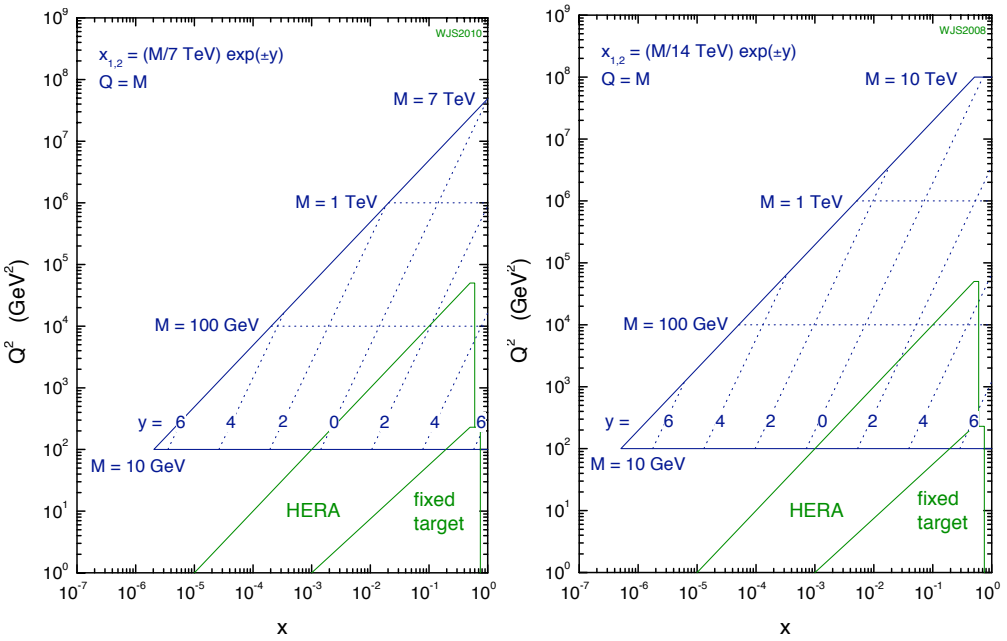

Figure 2.11: LHC **parton kinematics** - (x, Q^2)-plane for the production of a heavy system of invariant mass, M, and rapidity, y, at the LHC (in blue) and previous experiments (in green). For 7 TeV on the left and 14 TeV on the right [119]

2.4.2 Looking Forward

The central region is limited in its investigations of several processes such as phenomena at low-x, diffraction, luminosity determination with QED processes, and energy and particle flows for tuning cosmic ray event generators [120]. Moreover, studies have shown that a high fraction of the energy in collision events is at $\eta > 3$ [121] [122]. The increase in the available phase space (see Figure 2.11) and added information at large pseudorapidities (see Figure 2.12) necessitates measurements in the forward region. By extending the coverage of particle physics detectors through the addition of forward calorimeters on both sides of the experiment, measurements for higher η values become possible.

By opening up the available phase space for further emissions, the forward region is the perfect tool to study higher order reactions and small-x physics [123]. Alternative parton dynamics are expected to manifest themselves at such low-x regions, and studies on the full evolution become possible. Additionally, the forward region can distinguish between types of MPI (remnant-remnant, parton-remnant, or additional parton-parton);

investigations at large pseudorapidities could add information on colour reconnections of multi-parton interactions to beam remnants, the beam remnant itself, and possible long-range energy-multiplicity correlations based on trigger energy (or particle triggers) in forward detectors.

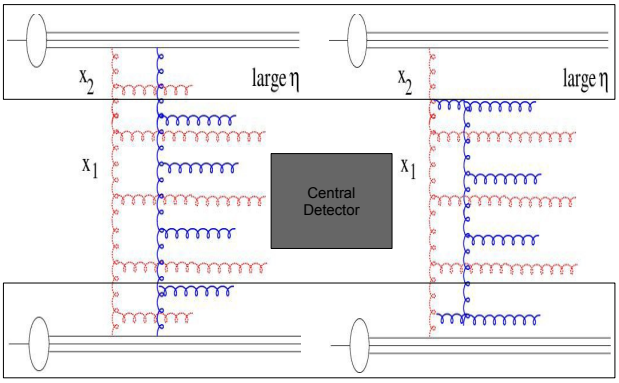

Figure 2.12: Forward region - Representation of the information gained through measurements at large pseudorapidities. The rectangles denote the forward region.

2.4.3 Energy Flow

Energy is a global variable that provides information on both neutral and charged particles. Energy flow is defined as $\frac{1}{N_{events}}\frac{dE}{d\eta}$, where E is the total energy of charged and neutral particles measured in bins of pseudorapidity. Among other things, the amount of energy flow generated by the additional partonic processes provides a way to distinguish between soft and hard MPIs [71]. If the transverse energy is too soft to form a jet within a kinematic region, it is a soft MPI. However, the energy produced may still be reconstructed and add to the jet, thereby creating a pedestal effect in the energy and E_T of the jets from the primary interaction. Semi-hard or hard secondary collisions lead to the direct formation of additional jets in the final state, leading to multi-jet events [60]. Therefore, at large pseudorapidities, energy flow is considered to be directly sensitive to the amount of parton radiation and multiple parton interactions.

2.4.3.1 Previous measurements

Previous measurements at proton-proton colliders found that the interactions were characterised by in increase in particle multiplicity dependent on $\sqrt{s}$, but a very small dependence of the particle transverse momenta on centre-of-mass energies [124], [125] [126]. A relation between the average multiplicity and transverse momentum of the produced particles was proffered [127], which described the measurements existing at the time. However, with higher luminosities in pp experiments, phenomena with an underlying dynamics involving hadron constituents became apparent [128].

The UA1 collaboration measured the transverse energy flow in proton-antiproton collisions over a large range of centre-of-mass energies (0.2 TeV – 0.9 TeV) at the SPS collider. They found that the average transverse momentum in an event was the same over a large variation of $\sqrt{s}$ for low particle multiplicity, but that it corresponded to the average p_T measured in events with a hard jet (high multiplicity events). That is, the evolution of the p_T distribution with increasing centre-of-mass energy implicitly lead to an increase in the average transverse momentum of the event. Large multiplicities and large transverse momentum were attributed to "gluon radiation from the interacting hadron constituents, or multiple low-x parton-parton interactions" [128]. The strong energy-dependent correlation between tranverse momentum and charged particle multiplicity was experimentally confirmed, and a parametrisation was proposed. They asserted that the average transverse energy in an event should increase with centre-of-masss energy, faster than $\ln(s)$ (see Chapter 8) [128].

Similar measurements were made at HERA, both in the traditional DIS framework and in the proton rest frame, over a wide kinematic range ($3.2\ \text{GeV}^2 < Q^2 < 2\,200\ \text{GeV}^2$, $8 \cdot 10^{-5} < x < 0.11$, $66\ \text{GeV} < W < 233\ \text{GeV}$) as a function of Q^2, x, W and pseudorapidity [32]. For the most part, the MC and data exhibited reasonable agreement with each other, over the kinematic range presented. However, the low-x behaviour ($\sim 10^{-3}$, i.e., near the proton remnant) could not be characterised by DGLAP-based models, as MC produced insufficient activity. Though BFKL calculations described the x-dependence better, the uncertainties were too large to draw any definitive conclusions on low-x parton dynamics.

Also, initial studies at HERA and measurements at hadron colliders failed to observe a significant dependence of the transverse energy flow on Q^2. Subsequent analysis of

data collected by the H1 collaboration provided the first experimental evidence for a rise in the transverse energy with Q^2 in the central pseudorapidity bin [32]. Other findings, such as a dependence on W and insensitivity to the nature of the colliding particles are consistent with data from hadron-collider experiments.

2.4.3.2 At the LHC

The LHC allows a similar study of the energy flow, for hadron-hadron collisions over a larger kinematic range and with a higher precision. Theoretical predictions of transverse energy flow as a function of pseudorapidity for minimum bias events and for events with a hard scale set by a central ($\eta < 2.5$) dijet system are presented in Figures 2.13–2.14. The distributions are generated with PYTHIA6.4 D6T at two centre of mass energies, $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, with and without multi-parton interactions, over the approximate η-range accessible to the Compact Muon Solenoid.

First and foremost, the plots (Figures 2.13– 2.14) illustrate the dependence of transverse energy flow on particle multiplicity; there is a large difference in magnitude between events with and without MPI-activity, often with twice as much transverse energy flow in a given pseudorapidity range. The flat distribution of E_T (plateau, in green) seen in Figure 2.13 indicates an even distribution of charged particles and activity throughout the range for events without MPI. The plateau ends abruptly, when there is very little energy left for additional activity; the edge of the plateau is related to the centre-of-mass energy, going further in η for higher beam energies ($\sqrt{s}$). Conversely, the distributions for multiple interactions in minimum-bias events (in black) fall, as there is less and less energy for particle production at larger pseudorapidities.

The presence of a hard, central dijet system ($p_{T,jet} > 8$ GeV and 20 GeV at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, respectively) produces the peaks observed in Figure 2.14. Following the hard scatter, there is little energy left for further particle production, which leads to the flattish slope in the non-MPI scenario. Both the presence of a dijet and the inclusion of multi-parton interactions increases the amount of radiation and particle multiplicity throughout the pseudorapidity range, a characteristic more pronounced by the difference in gradient within the range $3 < |\eta| < 6$.

The figures illustrate that significant differences in the distributions of transverse energy flow in the forward region ($|\eta| > 3$) can be observed through shape differences, i.e. beyond a simple difference in magnitude; the dissimilarity between the models are

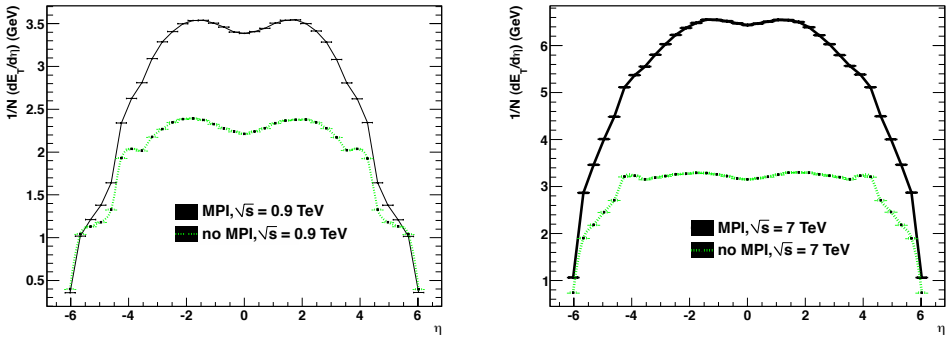

Figure 2.13: Transverse Energy Flow in Minimum Bias Events - Theoretical predictions of transverse energy flow in minimum bias events for $\sqrt{s}$ = 0.9 TeV (left) and $\sqrt{s}$ = 7 TeV (right), with (black) and without (green) multi-parton interactions, using PYTHIA6 Tune D6T. Note - different scales used for the two different centre-of-mass energies.

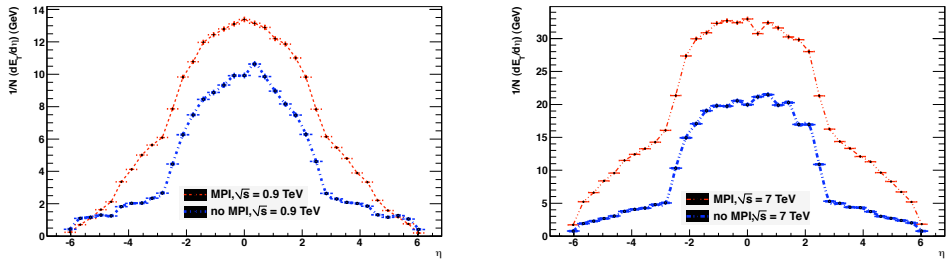

Figure 2.14: Transverse Energy Flow in Dijet Events - Theoretical predictions of transverse energy flow in dijet events for $\sqrt{s}$ = 0.9 TeV (left) and $\sqrt{s}$ = 7 TeV (right), with (red) and without (blue) multi-parton interactions, using PYTHIA6 Tune D6T. Note - different scales used for the two different centre-of-mass energies.

pronounced in the forward region. Distributions of energy flow display the same traits and expose the underlying physics expected - an increase in activity with respect to centre-of-mass energy, an η-dependence of the energy flow, and heightened activity in the forward regions (and therefore increased sensitivity to the underlying event) in the presence of a hard scale in the central region. Investigations using these observables provide complementary data to the existing UE studies, and go a long way towards improving our understanding of the underlying event and dependence on centre-of-mass energies ($\sqrt{s}$), in a previously unexplored phase space region.

3

The Experiment

3.1 Accelerator

3.1.1 Large Hadron Collider

Based near the Swiss-French border, and part of CERN (Conseil Européene pour la Recherche Nucléaire), the Large Hadron Collider (LHC) [129] accelerator collides protons (and heavy ions) with each other. Supplied by the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS), the LHC has run at three different energies since it became operational in 2009. Until the end of the 2009 running period, protons with energies of $E_p = 0.45$ TeV and 1.8 TeV were collided, generating a centre-of-mass energy, $\sqrt{s} = 0.9$ TeV and 2.36 TeV, respectively. In early 2010, the proton beam energy was increased to 3.5 TeV, generating a new centre-of-mass energy of 7 TeV. At present, the LHC has a centre-of-mass energy that is higher than any previous collider experiment, and is providing a substantial rate-of-collisions that will enable us to further probe the realm of high-energy physics.

The LHC has a circumference of 26.7 km. The protons travel in separate beam pipes, each of which can contain up to 1380 bunches in a single fill, separated temporally by 25 ns.

Six experiments use the LHC. There is one situated on each of the four straight sections that break the circular symmetry of the ring, as shown in Figure 3.1. The ATLAS [130] and CMS [131] experiments, which are located on the Swiss and French sides respectively, operate multi-purpose detectors to study the pp collisions. Both

detectors cover a solid angle approaching 4π and incorporate vertex, calorimeter and muon detectors.

The other two experiments with large detectors, ALICE [132] and LHCb [133] (shown in Figure 3.1), have more specific goals; the former is built to study heavy-ion collisions and heavy-quark physics, while the latter uses pp-collisions to study CP violation in b-meson decays. The remaining two experiments study forward physics in pp collisions; LHCf [134], based in the ATLAS cavern, and the TOTEM experiment [135], based in the CMS cavern.

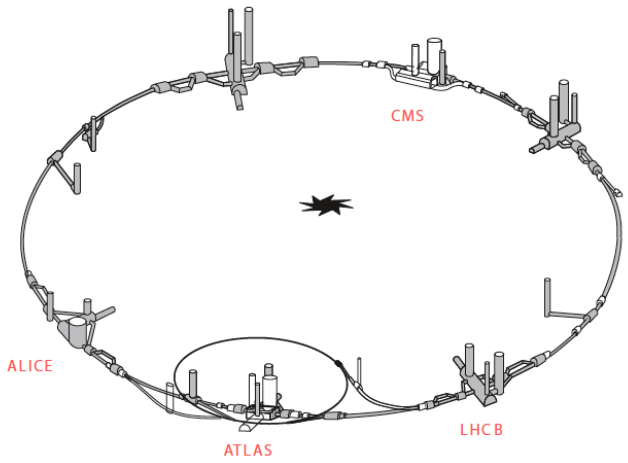

Figure 3.1: Graphical representation of the experiments at the LHC - Schematic layout of the LEP tunnel along with the four main experiments. ATLAS and CMS are the general purpose experiments, while ALICE explores heavy-ion physics in lead-lead collisions and LHCb investigates b-physics in proton-proton collisions [136].

3.2 Compact Muon Solenoid Detector

At present, the CMS collaboration comprises of 2000 physicists, from 155 institutes located in 37 different countries. A description of the detector built and currently used for data-taking is described herein. A more detailed description can be found in [131].

3.2.1 Coordinate System

CMS uses a right handed coordinate system, as shown in Figure 3.2. The z-axis points in the direction of the proton beam, and the x-axis towards the centre of the ring. In

the equivalent polar coordinates, r is the radial distance from the beam line, the polar angle, θ, is measured with respect to the z-axis and ϕ is the azimuthal angle. The pseudorapidity, η, is used to define the central and forward regions, in preference to θ.

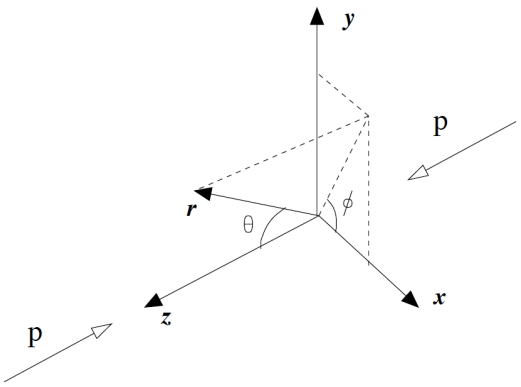

Figure 3.2: CMS **coordinate system** - Representation of the right-handed coordinate system used at CMS, including the beam directions.

3.2.2 Detector Overview

The Compact Muon Solenoid (CMS) detector is a general purpose detector designed to run at the highest luminosity at the LHC. It has been optimised for the search of the SM Higgs boson over a mass range from 90 GeV to 1 TeV, but it also allows detection of a wide range of possible signatures from alternative electroweak symmetry breaking mechanisms. CMS is also well adapted for the study of top, beauty and tau physics at lower luminosities and will cover several important aspects of the heavy ion physics programme. The experiment has chosen to identify and measure muons, photons and electrons with high precision. At the core of the CMS detector sits a large superconducting solenoid generating a uniform magnetic field of 4 T. The choice of a strong magnetic field leads to a compact design for the muon spectrometer without compromising the momentum resolution. The inner tracking system (cf. Section 3.2.3.1) will measure all high p_T charged tracks with a momentum precision of $\frac{\Delta(p)}{p} \sim 0.1\ p_T$ (p_T in TeV) in the range $\mid \eta \mid < 2.5$. A high resolution fully active scintillating crystals based electromagnetic calorimeter designed to detect the two photon decay of an intermediate mass Higgs, is located inside the coil (cf. Section 3.2.3.2). Hermetic hadronic

calorimeters surround the intersection region up to $|\eta| \approx 5$ allowing tagging of forward jets and measurement of missing transverse energy (cf. Section 3.2.3.3).

Figures 3.3 and 3.4 show, respectively, the CMS detector and the orientation of the various sub detectors from the side, as it was in the original design, and from the r-z plane. The detector as a whole, measures approximately 21.6 m in length, 14.6 m in diameter, and weighs around 12500 tons. It is almost hermetic, with only the beam pipes breaking the full coverage, and is asymmetrical, comprising of an additional detector (CASTOR) in the backwards direction. The main components of the detector are briefly described in the subsequent sections; namely, the inner tracking system, muon system, electromagnetic calorimeter and the hadronic calorimeter, the latter of which comprises of the Hadronic Forward Calorimeter (HF). The HF is described in detail, as it is the main component used in this analysis.

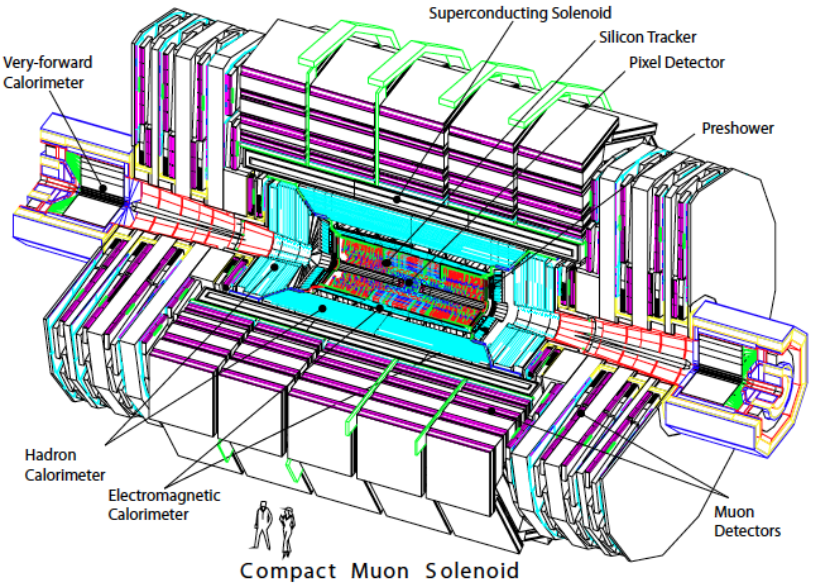

Figure 3.3: Perspective of CMS - CMS detector and its various components. [137]

3.2.3 Specific Components

3.2.3.1 Tracking System

The CMS tracker, with a length of 5.5 m and a radius of 1.25 m, is located around the interaction point, measuring charged particle trajectories and their production vertices.

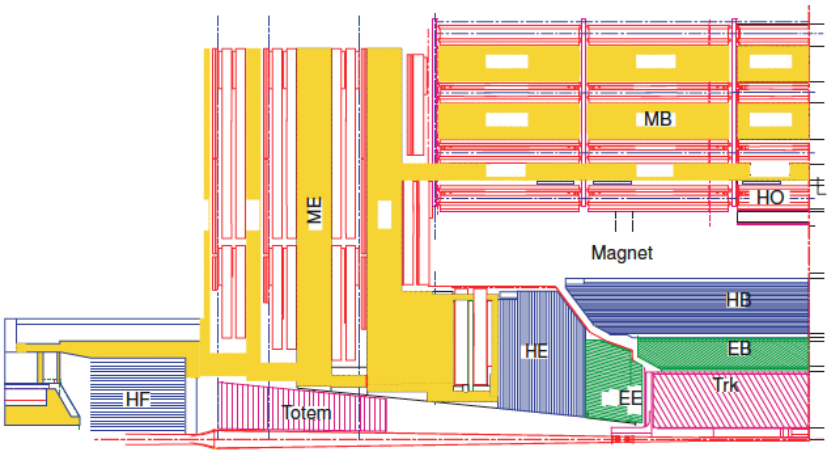

Figure 3.4: Quadrant of CMS in the r-z plane - The most important subsystems are shown in the r-z plane, i.e. the tracker (Trk.) inside the magnetic coil (Magnet) , the electromagnetic- barrel (EB) and endcap (EE) calorimeters, the hadronic barrel- (HB), hadronic endcap- (HE), outer hadronic- (HO) and hadronic forward- (HF) calorimeters, and finally the muon barrel (MB) and endcaps (ME) of the muon system. Additionally, the location of the tracking detector of the TOTEM experiment is shown. [138]

Situated within the solenoid, the magnetic field is at its strongest: $B = 3.8$ T along the z-axis. This field bends charged particle trajectories, in order to measure the momentum and charge from signals left in the detector by the particle.

An all-silicon detector, the tracker's material is radiation hard, allows a highly granular design with a fast response time. Owing to the high hit rate density close to the nominal vertex (see Table 3.1) and the desired occupancy of less than 1%, silicon pixel detectors are used in the innermost part of the tracker. Further from the nominal vertex, silicon micro-strip detectors with larger cell dimensions are used, with occupancies ranging from 2% to 3% per bunch crossing.

Table 3.1: Hit rate densities at LHC design luminosity.

Distance to nominal vertex	Hit rate density
4 cm	1000000 Hz/mm^2
22 cm	60000 Hz/mm^2
115 cm	3000 Hz/mm^2

There are three pixel layers, followed by ten strip layers in the barrel. Endcaps with silicon modules perpendicular to the beam axis complement the inner tracking system - two pixel disks and three inner plus nine outer strip disks on each side. The total pseudorapidity coverage is up to $\eta = 2.5$. Figure 3.5 shows the exact layout, as described here. This design meets the demands (high granularity, fast response) and accounts for the difficult operating conditions at the LHC.

The material budget varies from 0.4 X_o to 1.8 X_o, depending on the part of the tracking system. The majority of it is due to non-sensitive material such as mechanical support structures, cabling, cooling and electronics. This dead material producing multiple scattering is the largest limiting factor for the transverse momenta precision.

The construction and conception ensures good reconstruction of charged particles with transverse momenta above 1 GeV/c, within its volume (precisely). Measurements of secondary vertices and impact parameters allow efficient heavy flavour identification. Electrons are measured using a combination of information from the tracker and ECAL, while muons use the muon systems instead of the ECAL. Among other uses, track measurements provide an important input to high-level trigger algorithms.

3.2.3.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a composite, non-compensating calorimeter made of lead tungstate ($PbWO_4$) crystals, designed to measure energy deposits from electrons and photons. The use of such a material (high density, short radiation length, small Molière radius) results in a compact, fast, highly granular and radiation-resistant calorimeter, with good energy resolution. The composite nature of the calorimeter is depicted in Figure 3.6 and described below.

The barrel (EB) covers the pseudorapidity range $|\eta| < 1.479$, comprises of 61 200 crystals decomposed into 360 cells in azimuth and 2×85 cells in pseudorapidity, with signals read out using avalanche photodiodes. The total radiation length of the barrel crystals is 25.8 X_o.

The pseudorapidity range $1.479 < |\eta| < 3.0$ is covered by the endcap (EE), with one half (dee) on either side of the interaction point. Each dee has 3 662 crystals and the signals are read out using vacuum photodiodes. The length of the endcap crystals is 1.1 X_o shorter than for the EB, due to the presence of the preshower detectors in front of it.

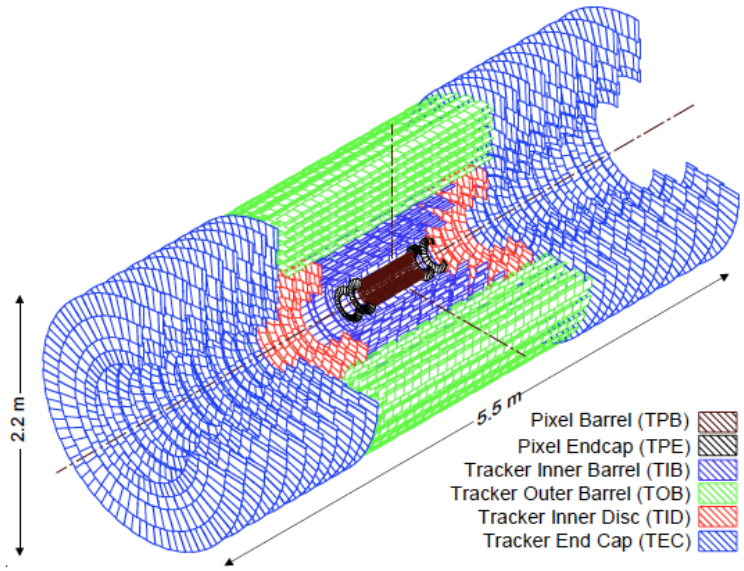

Figure 3.5: CMS Tracker and its components - The tracker is made up of a pixel detector and a silicon strip detector, each consisting of barrel layers in the central region and r-ϕ discs in the endcap region. The detector is installed around the beam pipe, covering up to $|\eta| = 2.5$. [138]

The preshower detector helps improve the position resolution of electromagnetic particles as well as identifying neutral pions, within the pseudorapidity range $1.653 < |\eta| < 2.6$. Silicon strip sensors measure the showers produced by particles interacting with the lead radiators.

The fine granularity of the ECAL leads to good shower position resolution and less shower overlaps. The scintillation decay time of the used crystals is relatively fast, where 80% of the light is emitted within 25 ns, thereby separating showers originating from different bunch crossings. Finally, the total depth of 25 X_o ensures the full-containment of electromagnetic energy, for energies below 500 GeV.

For energies above 500 GeV, shower leakage from the rear of the calorimeter is significant. In such cases, a simple parametrisation of the energy resolution (σ) is no longer valid. Disregarding such a scenario, the energy resolution of the detector is given by [139]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{E}\right)^2 \oplus \left(\frac{N}{E}\right)^2 \oplus C^2 \tag{3.1}$$

where S is a stochastic term (2.8%), N is a term for noise contributions (0.12), and C is a constant (0.3%). Each estimate has been confirmed with test beam data, using electron beams of energies 20 – 250 GeV/c.

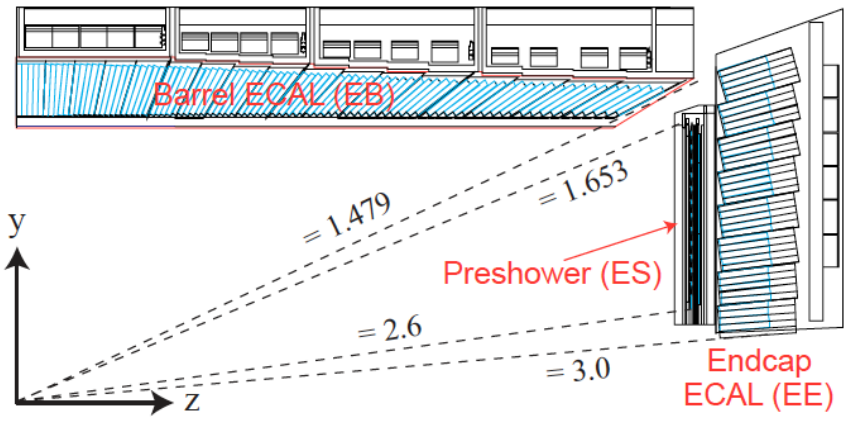

Figure 3.6: Longitudinal sketch of the Electromagnetic Calorimeter - Transverse section through the ECAL, showing the geometrical configuration of the electromagnetic barrel (EB), endcap calorimeter (EE) and the preshower detector (ES). [139]

3.2.3.3 Hadronic Calorimeter

The CMS Hadronic Calorimeter (HCAL) is a sampling calorimeter, providing coverage up to $\eta < 5$. Its depth of 10 interaction lengths prevents shower leakage, thereby preventing a degradation of the energy resolution. Moreover, the extensive coverage in pseudorapidity is essential for reliable measurements of missing transverse energy. Figure 3.7 is a longitudinal sketch of the HCAL, showing the geometrical configuration of its various components.

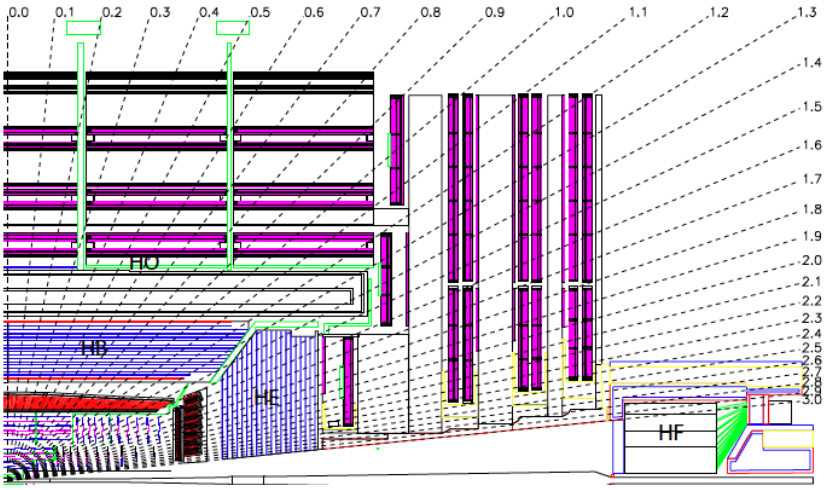

Figure 3.7: Longitudinal sketch of the Hadronic Calorimeter - Transverse section through the HCAL, showing the geometrical configuration of the hadronic barrel (HB), hadronic endcap (HE), hadronic outer (HO) and hadronic forward (HF) calorimeters. The numbers denote pseudorapidity. [137]

The HCAL barrel (HB) and endcap (HE) detectors are composed of alternating layers of brass (absorber) and plastic scintillators. Each scintillating tile is coupled to wavelength-shifting fibres, which are coupled to hybrid photodiodes. The detectors are divided into a tower geometry (see Figure 3.8). The HB covers $|\eta| < 1.3$ with each tower occupying 0.087×0.087 of η–ϕ space. The HE covers the pseudorapidity region $1.3 < |\eta| < 3$, with similar tower dimensions as the HB for $|\eta| < 1.6$ and $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for the more forward regions.

The Hadron Outer calorimeter (HO) is sandwiched between the solenoid and the muon system, to absorb energy leaking beyond the magnet. It provides further containment of hadronic showers (tail catcher), necessary for accurate missing-E_T mea-

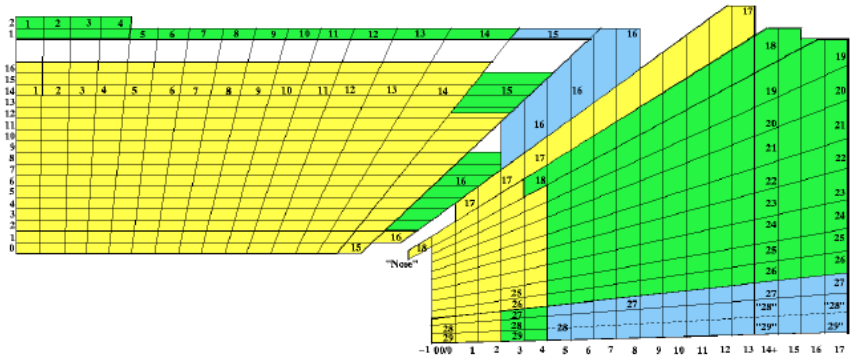

Figure 3.8: Tower structure of the CMS HCAL in the r-ϕ view - A schematic view of the tower mapping in r-z of the HCAL barrel and endcap regions. [139]

surements. It is composed of five rings, each 2.5 m long in z, and maintains the same scintillator geometry as the HB.

The Hadronic Forward calorimeter (HF), also a sub-system of the HCAL, is discussed in more detail later in this chapter (see Section 3.2.4.1), as it qualifies as a forward detector.

The energy resolution seen in beam tests for the HB, coupled with the corrections for the system's non-linearity, is:

$$\frac{\sigma}{E} = \frac{94.3\%}{\sqrt{E}} \oplus 8.4\% \tag{3.2}$$

with the energy given in GeV.

3.2.3.4 Muon System

The solenoid is surrounded by a dedicated muon detection and momentum measurement system. It uses three types of gaseous tracking chambers due to the large area, variations in particle flux and magnetic field, precision and speed; drift tubes (DT) in the barrel region ($|\eta| < 1.2$), cathode strip chambers (CSC) in the endcap region ($0.9 < |\eta| < 2.4$), and resistive plate chambers (RPC) in the barrel and endcap regions ($|\eta| < 1.6$) (see Figure 3.9).

In the barrel, where the magnetic field is less intense and the muon rate is low, high precision DT chambers are used. CSCs are employed in the endcaps as they cope well with the higher muon flux, neutron-induced background and the magnetic field. Due

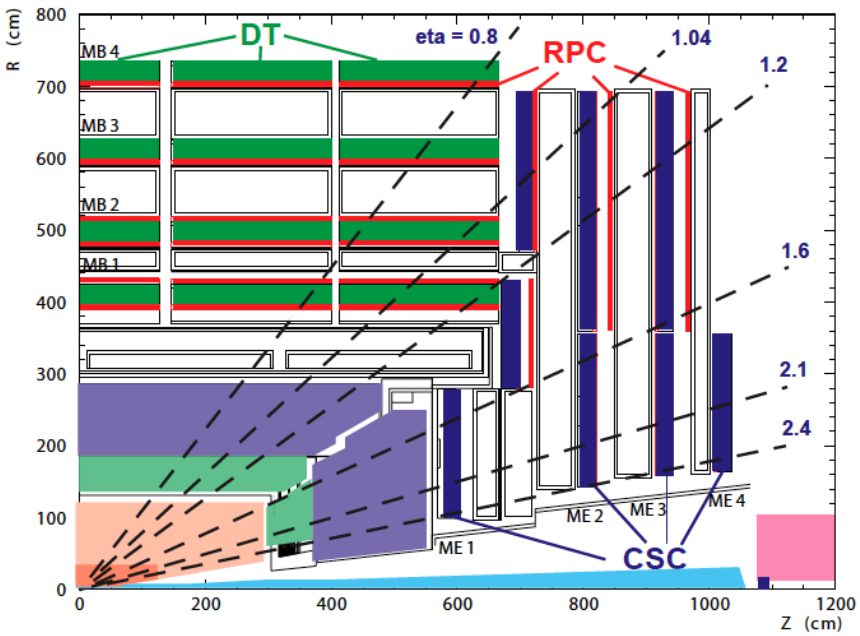

Figure 3.9: Layout of the CMS Muon System - . [131]

to their fast response with good time resolution, RPCs cover both regions, for triggering purposes and complimentary, albeit coarse position measurements.

In total, the system provides 25000 m^2 of active detection, with approximately 1M channels. A combination of the inner tracking system with the muon system provides the best muon momentum measurement, approximately 1% – 2% resolution for muons with $p_T = 100$ GeV/c .

3.2.4 Forward Calorimeters

3.2.4.1 Hadronic Forward calorimeter

The CMS Hadronic Forward (HF) calorimeters cover the range in pseudorapidity of $2.866 < |\eta| < 5.205$. They are located 11.2 meters from the interaction point on both sides of the detector and consist of iron absorbers (absorber) and embedded radiation hard quartz fibres (active material), which provide a fast collection of Čerenkov light. Half of the fibres run over the full depth of the absorber, while the other half start at a depth of 22 cm from the front of the detector (see Figure 3.11). These long (L) and short (S) fibres alternate in the HF structure, making it possible to distinguish between

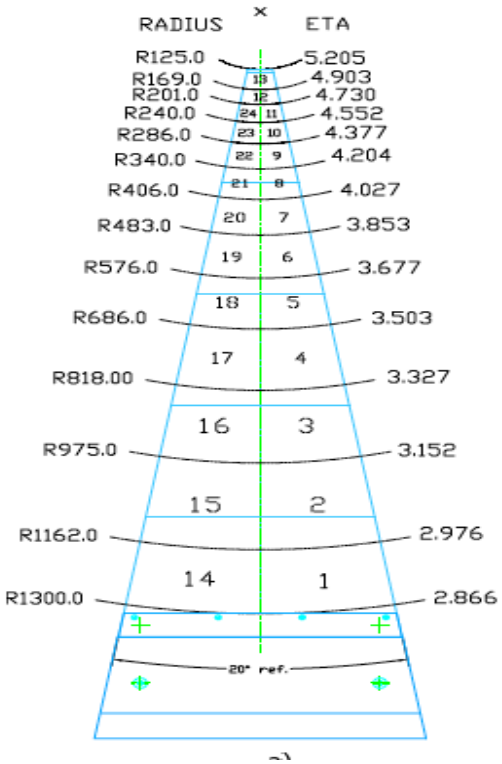

Figure 3.10: Transverse segmentation of the HF tower - [137]

showers generated by electrons and photons from those generated by hadrons due to differences in shower depths. The light from the fibres is detected by 8-stage photo-multiplier tubes (PMT) with borosilicate glass windows. They are housed in special read-out boxes. Each box contains 24 PMTs which corresponds to 10° in ϕ.

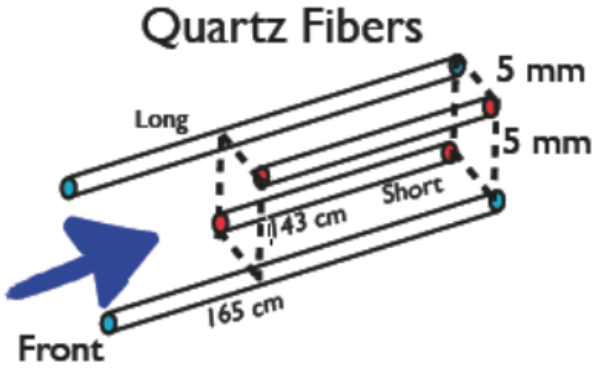

Figure 3.11: Long and short fibres in HF - Long and short fibres alternated to distinguish energy deposits of different particles types. Long and short fibres are separated in read out.

HF detectors use a ^{60}Co radioactive source for calibration. ^{60}Co produces two kinds of gamma rays with energies of 1.2 MeV and 1.3 MeV which produce Compton electrons. If the electron momenta are above the Čerenkov threshold, Čerenkov light inside quartz fibres are being created. The source is inserted into the centre of each readout tower by using a long wire and atomised source driver system. The charge difference in "source in" and "source out" histograms is used in three different methods (namely mean charge, fixed interval, and extrapolation methods) to calculate precision in calibration. These methods find the results for L and S fibre PMT readouts separately. Overall results on calibration precision vary between 5% and 10% for L fibre readout PMTs and from 6% to 20% for S fibre PMTs [140].

CMS uses a numbering convention for calorimeter towers, $i\eta$ and $i\phi$, which helps identify the towers internally; the HF has 27 segments in $i\eta$ and 18 segments in $i\phi$. These numbers are used for internal corrections and consistency checks (see Chapter 5.2.1.2, Chapter 7.1.5 and Chapter 7.2.7). Overall, the HF calorimeter is composed of 13 rings with a similar $\Delta\eta$ coverage of approximately 0.175 for each ring. The numbering convention of HF rings relevant to this analysis is given in Table 3.2. The energy flow

in HF is measured in corresponding bins of η, grouping two consecutive rings together, e.g., ring 2 and ring 3 ($4.903 < \eta < 4.552$), or ring 4 and ring 5 ($4.552 < \eta < 4.377$).

Table 3.2: HF ring numbering and η divisions.

Ring	$\Delta\eta$	η Boundary
1	0.302	5.205 - 4.903
2	0.173	4.903 - 4.730
3	0.178	4.730 - 4.552
4	0.175	4.552 - 4.377
5	0.173	4.377 - 4.204
6	0.177	4.204 - 4.027
7	0.174	4.027 - 3.853
8	0.176	3.853 - 3.677
9	0.174	3.677 - 3.503
10	0.175	3.503 - 3.327
11	0.175	3.327 - 3.152
12	0.176	3.152 - 2.976
13	0.110	2.976 - 2.866

The η segmentation of HF's geometry is done by using the radial distance from the centre of the detectors. As a result of this, all readout towers have $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$, whereas, as shown in Table 3.2, the actual segmentation of the first and last rings are much different: 0.111, and 0.300, respectively. These differences result in a disagreement of the energy response for these towers between data and event simulations. The disagreement arises from the difference in geometry implementation in the detector simulation for those two rings, with respect to the real detector. In ring 1, data show a higher response compared to the simulation, and in ring 13, data show a smaller response compared to the simulation. Issues with shower leakage, overlap between detectors (e.g. ECAL), and improper simulation by the shower library exaggerate the differences in energy response as well, particularly for the first and last rings. An update of the simulation is foreseen but will not be completed in time for this analysis. Therefore, in the present analysis, both rings (the innermost and the outermost) are excluded.

The region covered by ring 12 (see Table 3.2) is also covered by a section of the

electromagnetic endcap calorimeter (EE). In order to simplify the analysis and not consider the influence of the EE, ring 12 is also not used for the measurement. Taking all these considerations into account, the active area of HF is being restricted to $3.152 < |\eta| < 4.903$.

3.2.4.2 CASTOR

CASTOR is a tungsten-quartz Čerenkov calorimeter surrounding the beam pipe at 15 m to 16.5 m from the interaction point, just beyond HF at the end of CMS (Figure 3.12). It can be broken down further into two sections: the hadronic section which extends from $5.15 <| \eta |< 6.4 \sim 9.2$ interaction lengths (λ_I), and the electromagnetic section with a coverage of $5.3 <| \eta |< 6.5 \sim 22$ radiation lengths (X_0) [141].

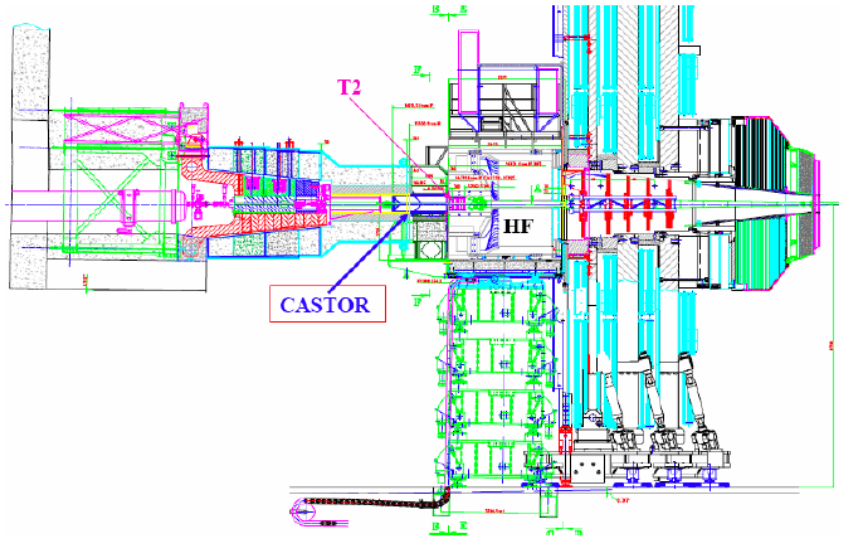

Figure 3.12: Location of CASTOR in the CMS forward region - [137]

Figure 3.13 shows the various components and the geometry of CASTOR. The calorimeter is made up of layers of tungsten plates (the absorber) and fused silica plates (the active medium). To maximise the Čerenkov light collected, the plates are at an angle of 45^o with respect to the particles flight direction. Light is produced by charged particles of the shower (like e^+e^-, protons, etc.) passing through the quartz. Internal reflections allow the light to reach the top of the silica plates which is then collected in the Reading Units (RU). Longitudinally, the full calorimeter has 2 EM sections and 12

hadronic sections, each of which constitute a RU used for detecting penetrating cascade particles. It is also symmetric in azimuth, divided into 16 ϕ-segments ($\phi = 22.5^o$) for better resolution. It is foreseen that additional positional resolution will be given by T2 from TOTEM. To avoid pile-up there is a need to take measurements during the low luminosity in the early LHC phase.

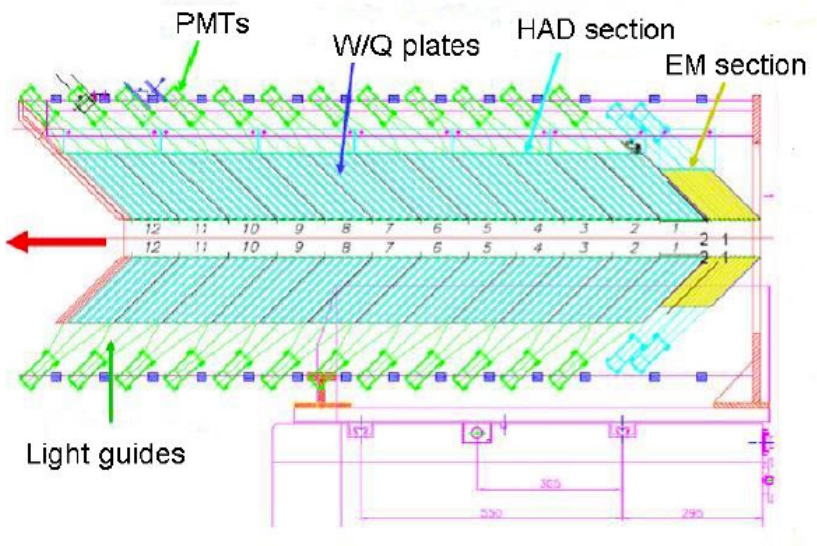

Figure 3.13: Details of the components and geometry of the CASTOR calorimeter. - [137]

3.2.5 Trigger and Data Acquisition

A trigger is essential as the bunch crossing rate at the LHC is $\sim$ 31.6 MHz, yielding $\mathcal{O}(10^9)$ events per second; a rate which is far too fast, providing more data than can be written to tape. Since only $\mathcal{O}(10^2)$ events per second can be archived, CMS reduces the rate using a two-level trigger system. This architecture forms part of the data acquisition (DAQ) chain to select the *pp* events. The trigger is designed to ensure that this is done with as high an efficiency as possible within the limitations of the detector. Along with the trigger system, the CMS DAQ system also collects and analyses signals from CMS front-end electronics at the reduced trigger rate.

The Level-1 Trigger (L1T) is hardware based, but custom-designed with flexible logic and programmable electronics. Designed to reduce the rate to below 10 kHz,

there are individual component L1Ts which consider information from the calorimeters and muon systems to search for interesting physics objects (electrons, muons, jets, etc). These local triggers are linked via the Global L1T. The decision whether to keep an event or not takes longer than the time between the bunch crossings and so the data are buffered in a 3.2 μs pipeline, after which time the on-detector buffers lose the event under consideration. Based on the information available within the 3.2 μs (1 μs computation, 2.2 μs cable latency) allocated to make the decision, an event is selected or rejected. The architecture of the L1 trigger is shown in Figure 3.14.

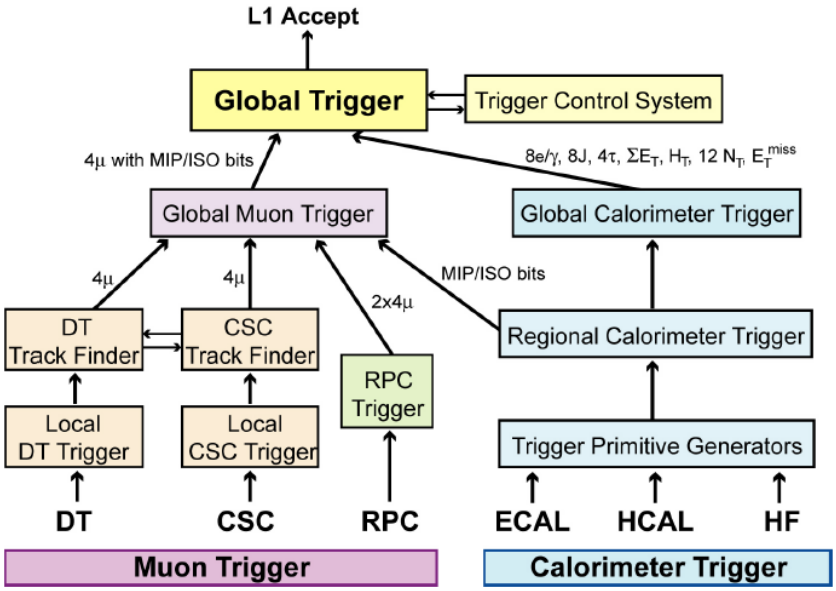

Figure 3.14: Architecture of the L1 trigger - Trigger path from the separate CMS sub-detectors like the ECAL or CSC to the global L1 trigger. [137]

Following a Level-1 accept, the data are sent to the High Level Trigger (HLT) which further reduces the output rate from 100 kHz to 100 Hz for offline analysis. The HLT is a software trigger, with a processor farm and a switching network designed for this purpose. All sub-detector information is available to the HLT for it to make its decision. In fact, the information available at the HLT is less crude than the L1T, as it includes quantities based on limited charged particle tracking, calorimeter timing, jet finding, the vertex position and various other kinematic quantities.

Each physics group has a set of HLT paths, which, using the HLT quantities, selects events relevant to the physics studied within that group. One or a combination of

paths need to be selected for an event to pass the HLT. Each time the HLT sends an accept signal, the event is classified according to the paths activated and written to tape. Together with the event data, the information is passed onto the HLT farm. Finally, the events are reconstructed offline, and pass onto the next stage: distribution, re-reconstruction, analysis.

Since the computing needs of such an experiment exceed the capabilities of a single site, computing resources are combined into a common hierarchical structure of Tiers (see Figure 3.15. The CERN Analysis Facility (CAF) is a Tier-0 centre, where the experimental data is recorded and an initial reconstruction of data is performed, before distribution to the Tier-1 centres.

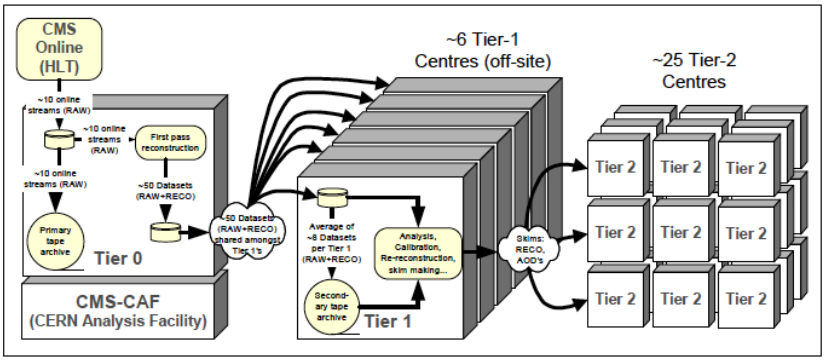

Figure 3.15: Data flow in CMS - Flow of data from the experiment and first processing (Tier 0) to the analysers (Tier2). [137]

The six off-site (i.e., not at CERN) Tier-1 centres receive the raw and reconstructed data for additional mass storage, a re-reconstruction, and data-intensive central analyses. These are large computing farms with fast links amongst themselves and to the T-0 centre. Finally, data is distributed amongst twenty-five Tier-2 centres, ready for user-analyses. The bulk of user-related tasks such as simulation studies, offline calibration, and data analysis are performed at these sites.

CMS is a computing-intensive experiment, requiring a multi-tiered, world-wide computing model.

4

Event Simulation and Reconstruction

4.1 Event Simulation

Event simulation plays a key role in many aspects of HEP, from calibration of detectors to predictions of physics signals and their backgrounds. In this chapter, an overview of such simulations is presented.

4.1.1 Monte Carlo Method

Monte Carlo methods are stochastic techniques. That is, the term applies to any algorithm developed to compute results through repeated random sampling. Such a method is particularly useful when analytical calculations are unfeasible (e.g. fixed-order calculations at NLO), and for simulating systems with many coupled degrees of freedom (e.g. high-dimensional cases like molecular dynamics). Often, it is also used to complement theoretical calculations, when an analytical solution is viable.

While the exact implementation varies according to the system chosen, the general algorithm is as follows [142]:

- set a range of possible inputs;
- generate inputs through (pseudo-) random sampling of the range;
- perform a calculation of the function relevant to the system;

- collate the results.

The results are reliable as long as the inputs are truly random (or large enough to appear random w.r.t. the domain), and there is a large number of inputs.

In particle physics, certain situations cannot be solved analytically but a numerical approach is valid; for example, modelling the complex final states in hadron-hadron collisions with high particle multiplicities, in order to correct for detector effects and compare measurements to theoretical predictions. Based on the kinematics of an event, random numbers can choose a hard process from the available matrix elements; after choosing a process for the event, the generator proceeds to simulate that interaction. By generating a large sample, we obtain statistically accurate and relevant predictions such as cross sections (probability distributions). Thus, using the Monte Carlo method as the underlying technique, event generators are employed to simulate collision events. The various effects considered in such a simulation, such as parton evolution and hadronisation, have been discussed in Chapter 2.3. The manner in which those effects are linked through factorised steps is common to most generators and is presented in the next section.

4.1.2 Event Generation

The first step is to calculate the matrix element at a fixed order of the strong coupling (α_s), with the momenta of the in-going partons being randomly chosen based on the input parton distribution functions (PDFs). Given the probabilistic nature of physics at this scale, the initial conditions of the in-going partons do not produce the same final state for each event; the momenta of the outgoing partons is randomly distributed in the available phase space.

Following the calculation of the hard process, higher order QCD effects are added using parton shower models. The evolution of the partons are governed by different evolution equations, depending on the generator used (see Chapter 2.3).

Next comes the hadronisation of the partons. Tuned to data, hadronisation models are exclusively phenomenological. They cannot be calculated in perturbation theory as this effect occurs at energy scales $\approx \Lambda_{QCD}$. The most widely used event generators often differ in their choice of models. At this stage, the modelling of MPIs and the

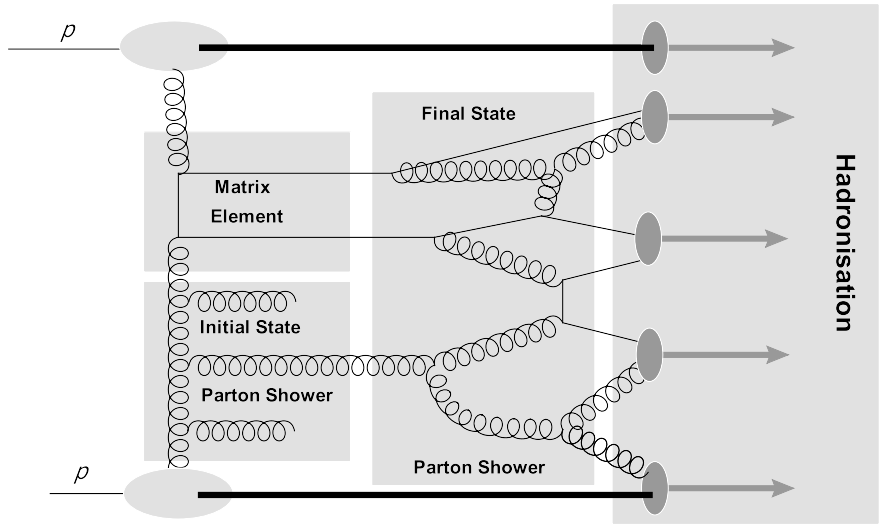

Figure 4.1: Steps in event generation - The basic steps common to most event generators. Concept taken from [143].

underlying event are also taken into account, as they depend heavily on the description of hadronisation. Finally, the decay of unstable hadrons to stable particles is simulated.

The output or final quantities obtained are called hadron-level or generator-level quantities. Information from this level is often used to calibrate the detector response, as well as tuned to data to reduce deviations.

4.1.3 Detector Simulation with GEANT4

GEANT4 (GEometry ANd Tracking) is a MC-based simulation toolkit which links hadron-level information to data collected at experiments, by mimicking detector response and output signals [144]. Through the accurate simulation of particle interactions with various materials and detector geometry, results from event generators become comparable to actual data.

CMS software (CMSSW) [145] uses the GEANT4 simulation toolkit [144] to simulate detector effects. Not only are all physical regions (detector, support structure, etc.) of the detector simulated, but the magnetic field's effects on the detector response are also modelled (see Chapter 6) [146]. Up-to-date information on dead channels and poorly calibrated components are stored in a 'conditions' database, which are used to produce realistic event samples on a regular basis.

Particles from the final state produced by the event generators are decayed according to their branching ratios and kinematics. As these particles pass through the detector, the interaction of the particles with the detector material is computed. Detector noise, channel cross-talk, and digitalisation steps applied by each sub-detector are also included in the simulation. Finally, a full emulation of the L1T and HLTs is added.

The interaction of particles with the material is modelled and parametrised using test beam data for the individual subsystems. Further deviations between data and simulation are minimised by tuning the simulations to collision data. However, some differences still remain, as there are many challenges in modelling phenomena such as hadron cascades.

4.2 Event Reconstruction through Physics Object Reconstruction

The output from data streams and from the simulation-digitisation step have the same structure and format. Calibration and conversion constants are stored in databases, updated regularly, to ensure similar detector response after the digitisation step. Therefore, the same reconstruction algorithms are used for both the simulated samples and collected data; the analyses are independent of the provenance of the sample.

4.2.1 Physics Object Reconstruction

Most physics analyses involve basic physics objects such as electrons, photons, muons and jets. These objects define the final state of a hard scattering process.

4.2.2 RecHits and CaloTowers

The ECAL has been calibrated to electrons depositing energy in an array of 7×7 crystals, while the HCAL has been calibrated to 50 GeV pions for 3×3 (4×4) cells in the HB (HE). Representing a certain amount of energy in a given calorimeter channel (no threshold, pedestal subtracted), RecHits are reconstructed objects of a calorimeter composed of such cell clusters.

CaloTowers are the result of combining ECAL and HCAL RecHits into towers, limited by the granularity and geometrical structure of the HCAL. They are combined according to a threshold scheme.

4.2.3 Particle Flow Reconstruction

The concept of particle flow (PF) was originally introduced at ALEPH [147] and employed at hadron colliders in a limited way [148]. At CMS, it is used as a standard reconstruction method, producing physics objects widely used throughout the collaboration. The idea is to drawn on information from all the sub-detectors in order to reconstruct a coherent physics object and improve the overall measurement resolution. The result is a list of candidates (PFCandidates); electrons, photons, muons, taus, charged and neutral hadrons, with properties of the particle they represent and accurate four-momenta [149].

The particle flow method seeks to reconstruct the event in terms of individual particles as well as create the basis for composite objects such as jets. The high granularity and excellent momentum resolution of the CMS tracker offsets the limited energy resolution of the calorimeters. The essential building blocks of the algorithm are tracks, clusters and muons. Figure 4.2 gives an overview of how these building blocks are combined, using information from all the sub-detectors, to create PFCandidates.

CMS uses an iterative tracking algorithm to convert TrackRecHits into high purity tracks with a negligible fake rate [151]. With a low contamination rate ($\approx 1\%$), particles with momentum as low as 150 GeV/c can be reconstructed. Tracks from muons are combined with signals from the muon system, forming the Global Muon collection. The remaining reconstructed particles form the PFTrack collection.

In a manner similar to the tracks, a clustering algorithm converts calorimeter RecHits to PFClusters. RecHits are re-classified as PFRecHits by adding information on their spatial coordinates (η, ϕ and x, y, z). This new collection is used as input to the PF clustering algorithm, which produces PFClusters with clustering parameters specific to each sub-detector. Provided the showers do not overlap, "clustering"[150] separates energy depositions that are close together; it distinguishes between charged and neutral particles.

Given that particles generally interact with more than one sub-detector, it is important to link the elements from the various sub-detectors if they originate from the same particle. Linking uses the "block algorithm" [150] to consider all the elements in

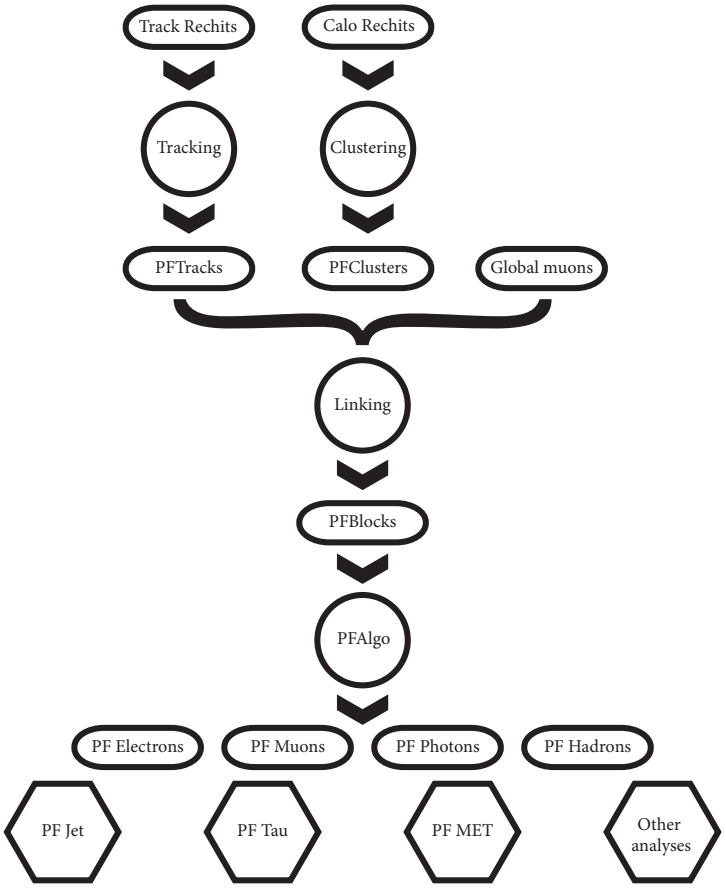

Figure 4.2: Flowchart for Particle Flow reconstruction - An overview of the algorithm used to reconstruct particle flow objects [150].

the event and link them based on their spatial distributions, creating PFBlocks. If a PFTrack cannot be connected to a PFCluster, and distribution of the HCAL hits overlaps or envelopes the ECAL ones, the calorimeter cells are associated to each other instead.

Finally, the PF algorithms produce the individual particle lists (PFCandidates) by analysing the content of the PFBlocks. These lists form the basis for composite objects such as jets or missing energy. More information on the exact algorithm and its performance is described elsewhere [149] [150].

Due to the absence of tracking information, particle flow only uses calorimeter information at large pseudorapidities. Therefore, particle flow objects are disregarded for the measurement of the energy flow with the HF. However, the list of PF particles is used as input for the jet type (PFJets) used in this analysis.

4.2.4 Jet Reconstruction

Jets are narrow cones of hadrons and other particles produced by the hadronisation of quarks or gluons. These bundles of hadrons are central to most analyses as they contain a lot of information pertinent to the hard collision; they cut back to the hard partons emerging from the hard scattering, providing access to QCD dynamics. Therefore, to probe QCD matter created in a collision event, good jet reconstruction is essential.

4.2.4.1 Jet Algorithms

The purpose of jet algorithms at collider experiments is to reduce the complexity of the final state, by simplifying the multitude of hadrons to simpler objects; they map the momenta of the final state particles into the momenta of a certain number of jets [152]. The extension of the jet is controlled by a resolution or distance parameter, R, producing well defined physical objects. Running a jet definition produces a physical observable which can be measured. Requirements of such an algorithm are that it be simple to implement in an experimental analysis (speed) and the addition of a soft particle or a collinear splitting should not change the final hard jets (infrared and collinear safety, IRC) [153].

There are two main classes of jet algorithms, sequential recombination algorithms [154] [155] and cone algorithms [156]. Sequential recombination algorithms have a bottom-up approach, where they combine particles starting from those closest to the initial seed or input. This involves setting a distance criterion and iterating the recombination until

few input objects are left. The resulting object is called a jet. Cone algorithms have a top-down approach which consists of finding coarse regions of energy flow. A stable cone is one where its axis coincides with the sum of momenta of particles in it, forming a jet.

The Iterative Cone algorithm is one of the first cone algorithms developed, often used due to its speed and efficiency; at CMS, its widespread implementation during the early days of the experiment makes it the default choice for many tasks. However, it is not IR-safe. The SISCone is an improvement on the standard iterative cone algorithms, made faster and IRC safe by inventing circular enclosures to find stable cones (given radius, R, in η–ϕ) [156] [157]. Due to its speed and predictable runtime, the IC-jets are used for triggering purposes at CMS [158].

The k_t algorithm is a recombination algorithm which work as follows [154] [155] [159]:

- calculate the distances between a pair of particles, d_{ij}, as defined in Eq. 4.1;
- calculate the beam distances, d_{iB}, as defined in Eq. 4.2;
- combine particles with smallest distance or, if d_{iB} is smallest, call it a jet, removing particle i from the list;
- find the smallest distance again and repeat procedure until no particles are left.

$$d_{ij} = min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta y^2 + \Delta \phi^2}{R^2} \tag{4.1}$$

$$d_{iB} = k_{ti}^{2p} \tag{4.2}$$

Here, k_{ti}^2 is the transverse momentum, R is a jet-radius parameter to weigh the distances d_{ij}, and p sets the power of the momentum scale. The behaviour of the algorithm depends strongly on p, the free parameter of this algorithm. The three commonly used and well-known cases are the k_T algorithm ($p = 1$), the Cambridge-Aachen algorithm ($p = 0$) and the anti-k_T algorithm ($p = -1$). Anti-k_T is the default choice for most physics analyses at CMS due to IRC-safety, speed[1] and stability of the algorithm [161]. Another reason for anti-k_T being the preferred jet algorithm is due to its shape; anti-k_T jets have a more regular, conic form, compared to the other k_T algorithms.

[1]The k_T-scheme was improved in speed by reducing the problem to near-neighbour searches [160].

4.2.4.2 Jet Types

For each of the previous jet definitions, the smallest seed considered was a particle. With generator-level simulations, the use of particles to produce generator-level jets or GenJets is possible. In experimental data, reconstructed objects are the input for jet algorithms. There are four different types at CMS, based on the four combinations of reconstructed objects that can be employed.

The first approach is to use the energy deposits in the calorimeters, using CaloTowers to produce CaloJets. The poor energy resolution of the HCAL limits this jet-type from being effective for most analysis purposes. Taking advantage of the excellent momentum resolution of the tracking system, tracks are matched to the CaloJets, thereby improving on the energy measurement. This constitutes the second approach, forming jet-plus-track (JPT) jets. Creating jets solely from the tracks as input is the third method. However, the energy scale of these TrackJets is comparable to CaloJets, rendering them insufficient for physics analyses. Their utility lies in serving as a cross-check for the other jet types. The fourth and final way is to profit from the efficacy of PFCandidates (see Section 4.2.3), exploiting the performance of all the sub-detectors, building PFJets from particle flow objects.

4.2.4.3 Jet Energy Correction

The reconstructed jets (RecoJets) are calibrated in a set of steps, shown in Figure 4.3 [162].

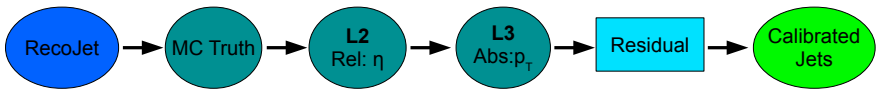

Figure 4.3: Flowchart for jet energy corrections - The set of steps taken to correct reconstructed jets to calibrated jets.

The first step is to correct reconstructed jets by matching them to generator-level jets in simulated events. The ratio of the two collections as a function of p_T and η supply a global calibration factor, which is subsequently applied to RecoJets from both data and detector-level simulations. Discrepancies between detector-level simulations and the real data result in additional η- and p_T-dependent corrections, also called relative (L2) and absolute (L3) corrections respectively.

L2 or relative corrections ensure a homogenous, flat energy response in η. Jets with higher values of absolute η are adjusted with respect to jets in the central region, by employing momentum conversation (p_T balance) in back-to-back dijet events. L3 or absolute corrections scale the p_T response of the jets. γ+jet events provide the corrections by comparing the recoil of the jet to the photon measured in the ECAL. The poor HCAL resolution disfavours the use of other (hadronic) event types for L2 corrections. While L2 and L3 factors can be derived in a data-driven way, simulated events were used in this analysis. Consequently, deviations between data and simulations remain. Additional adjustments to p_T and η due to such residual differences at high $|\eta|$ are applied to data [163]. The combination of simulation-based and residual corrections reduces the total systematic uncertainty on the jet energy scale.

This analysis uses anti-k_T PFJets with a size parameter $R = 0.5$, using the correction scheme shown in Figure 4.3.

5

Event Selection & Measurement

5.1 Event Selection

5.1.1 Data Collection

Between November 2009 and August 2011, the LHC delivered an integrated luminosity of 1.74 fb^{-1}. Runs in which all of the central components were functional and conditions were good for data taking are included in the Run Registry. These data are also required to pass a number of quality assurance tests. The amount of data recorded over the last few months is shown graphically in Figure 5.1, alongside the total integrated luminosity provided by the LHC. The majority of these data were not included in this analysis, as the instantaneous luminosity during the latter running period was significantly higher. Their inclusion would be non-trivial and would introduce additional systematic effects. Moreover, the statistical errors on data used is low, and so the improvement by the inclusion of further runs is not warranted.

5.1.2 Online Trigger Selection

For the trigger selection, several HLT paths are used on data. The CMS trigger system [131, 165] was used together with two elements of the CMS detector monitoring system, the Beam Scintillation Counters (BSC) [166] and the Beam Pick-up-Timing for the eXperiments (BPTX) devices [167]. The two BSC detectors are located at a distance of 10.86 m from the nominal interaction point, in both directions along the beam axis, covering the $|\eta|$ range from 3.23 to 4.65. Each BSC consists of a set of 16 scintillator

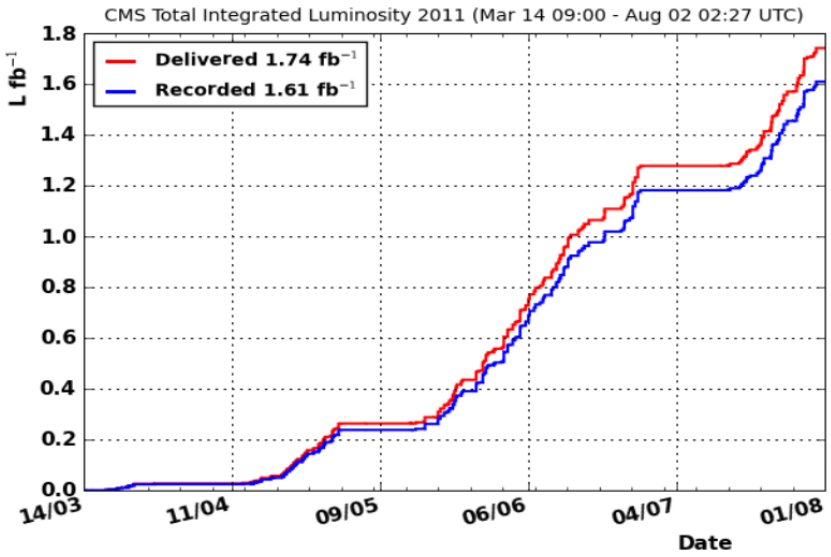

Figure 5.1: Data delivered by the LHC - Total integrated luminosity delivered by the LHC (in red) and data recorded by CMS, till August 2011. Plot taken from [164].

tiles. They provide information on hits and coincidence signals with an average detection efficiency of 96.3% for minimum ionising particles and a time resolution of 3 ns, compared to a minimum inter-bunch spacing of 25 ns for data. Located around the beam pipe at a distance of 175 m from the IP, the two BPTX devices are designed to provide precise information on the structure and timing of the LHC beams, with a time resolution better than 0.2 ns. The various components of the minimum bias trigger used in this analysis are:

- the beam-crossing trigger to ensure collisions;
- the trigger based on coincidence of the Beam Scintillation Counters on both sides, to reject single diffractive events;
- a trigger to veto beam halo events, in order to reduce beam backgrounds and prevent contamination of the data samples;
- a flag signalling good data taking conditions, i.e., stable beams from the LHC and key detector components ready for data-taking (e.g. tracker).

5.1.3 Offline Event Selection

Events which passed the trigger selection were subjected to further selection requirements offline, designed to clean data. In order to reject events with interactions between beam protons and rest gas in the beam pipe (beam scraping events), a minimum of 25% high purity tracks, i.e. tracks which pass tight selection criteria (details of selection criteria in [168]) is required if there are more than ten tracks in the event. Also events with a significant amount of noise in HCAL are rejected. Finally, the presence of at least one good primary vertex is necessary to ensure a collision occurred.

The vertex must fulfil certain quality criteria outlined by CMS; proximity to the nominal interaction point, with a limit on the number of degrees of freedom (NDOF) for the vertex fit (χ^2). The number of degrees of freedom parametrises the quality of the vertex, and is given by:

$$NDOF = 2 * \sum (weights) - 3 \tag{5.1}$$

For the vertex selection algorithm currently used by CMS, NDOF is related to the number of tracks associated to a vertex ($weights = 1$ for perfect tracks) [169]. In this analysis, the reconstructed position of the vertex along the beam line must be within ± 15 cm of the nominal CMS detector centre ($|z| < 15$), and the distance in the transverse plane from the nominal beam line within a 2 cm radius ($r < 2$ cm), with at least five tracks associated to it (NDOF ≥ 4).

With each requirement comes the risk of introducing inefficiencies and increasing the systematic uncertainties. Additionally, they can bias the phase space, thereby affecting the final cross section definition and complicating the interpretation of the measurements. However, these quality criteria are strict enough to efficiently clean data and reduce beam gas contamination from the event samples while minimising their impact on the total minimum bias cross section [168] [170].

Figure 5.2 gives an estimate of the number of events rejected by each condition.

5.1.4 Signature of Events

It is the aim of this analysis to study energy flow for two different event topologies; minimum bias events and events with a central dijet system. The presence of a hard scale set by the dijet system impacts the parton evolution within an event, thereby

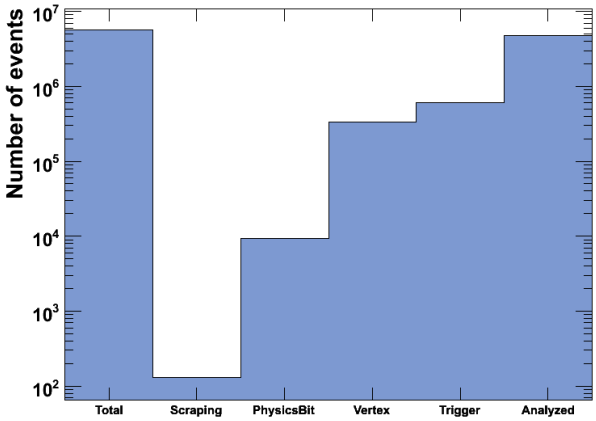

Figure 5.2: Events remaining after selection - The first bin gives us the total number of events before selection for the $\sqrt{s} = 900$ GeV minimum bias sample. The following bins show the number of events rejected by each condition, while in the last bin, the number of events analysed is presented. [171] [172]

requiring different dynamics of the proton's hadronic constituents from that in a minimum bias event; dijet systems are expected to be more sensitive to QCD phenomena. Figure 5.3 depicts the different event topologies, using the CMS event display software.

The comparison of the energy flow for these two event classes, minimum bias (zero or more partonic interactions) and dijet events (one or more partonic interactions), allows us to differentiate between soft and hard QCD contributions to the underlying event activity and investigate the validity of evolution models in a previously unexplored phase space.

The phase space being examined here is intended to be complimentary to that of previous energy flow measurements, and to gain sensitivity to the underlying event. The overriding factor when defining the kinematic region is the limitation of the detector (see 3.2.4.1).

5.1.5 Di-jet Event Selection Criteria

The dijet selection criteria can be broken down into two components: kinematic phase restrictions and jet quality criteria. The phase space requirements are that the two jets with the highest transverse momentum be in a central region ($\eta < 2.5$), back-to-back ($\Delta\phi - \pi < 1.0$), with a p_T threshold scaled in proportion to the centre-of-mass energy.

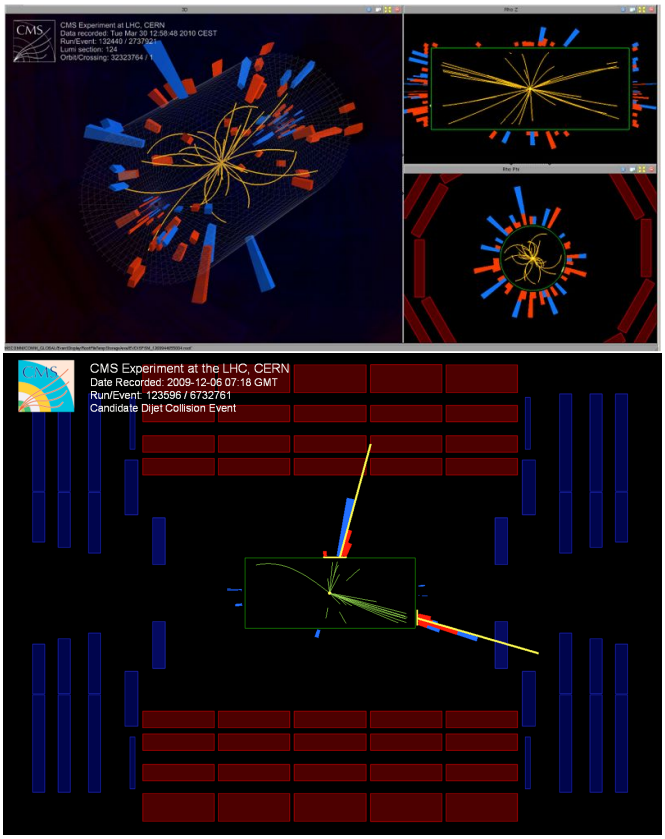

Figure 5.3: Event displays for minimum bias and dijet events - Top: CMS event display for a random event that passed the minimum bias selection criteria. Bottom: A dijet event in RZ-view. Taken from [173].

PFJets (see Chapter 4.2.4) are well-defined within the specified pseudorapidity range, and the η-restriction prevents jet energy leaking into the HF which would contaminate the energy flow measurement.

The back-to-back condition is aimed at reducing the angular decorrelations in the event, i.e. limiting parton radiation. In general, the requirement is applied in certain UE studies, in order to reduce the contribution of radiation (parton showers) to the underlying event, thereby allowing a little more sensitivity to MPI effects. Conversely, a small $\Delta\phi$ is heightens sensitivity to parton showering. Preliminary investigations show that the criteria does impact the overall forward energy flow but the effect is minimal.

The lower limit for jet p_T is designed to ensure that the two hardest jets are relatively well constructed. Moreover, it is statistically advantageous to keep the momentum threshold low, given the size of the data sample and scaling of multi-jet events. The requirement on the minimum transverse momentum was chosen in order to give comparable limits on the fractional momentum (x_{min}) carried by the jets for the two centre-of-mass energies; $x_{min} = \frac{2p_T}{\sqrt{s}} \approx 1.8 \cdot 10^{-3}$ $(0.6 \cdot 10^{-3})$ for $\sqrt{s} = 0.9$ TeV (7 TeV).

Table 5.1: Selection of anti-k_T dijets for data and simulated samples at $\sqrt{s} = 900$ GeV. [174]

Variable	Selection
p_T	> 8 GeV
η	< 2.5
$\Delta\phi - \pi$	< 1.0
Neutral hadronic fraction	< 1.0
Neutral electromagnetic fraction	< 1.0
Number of constituents	> 0
Charged hadronic fraction (if jet $\eta < 2.4$)	> 0.0
Charged electromagnetic fraction (if jet $\eta < 2.4$)	< 1.0
Charged multiplicity (if jet $\eta < 2.4$)	> 0

Jet quality criteria (JetID) include limits on the fraction of energy measured in the ECAL (electromagnetic fraction) and HCAL (hadronic fraction) from both charged and neutral particles; PFJets should contain at least one charged particle, if they are within the tracker limit. These values ensure generic noise-rejection and low fake-jet rates while maintaining high efficiencies. A summary of the requirements is listed in

Table 5.2: Selection of anti-k_T dijets for 7 TeV. [175]

Variable	Selection
p_T	> 20 GeV
η	< 2.5
$\Delta\phi - \pi$	< 1.0
Neutral hadronic fraction	< 0.9
Neutral electromagnetic fraction	< 0.9
Number of constituents	> 0
Charged hadronic fraction (if jet $\eta < 2.4$)	> 0.0
Charged electromagnetic fraction (if jet $\eta < 2.4$)	< 1.0
Charged multiplicity (if jet $\eta < 2.4$)	> 0

Table: 5.1 (5.2) for $\sqrt{s} = 0.9$ TeV (7 TeV). The values listed are the recommended jet quality criteria, as they result in negligible biases at the analysis level [174] [175].

5.2 Measurement

5.2.1 Technical Issues

5.2.1.1 HF Fibre Response

The HF calorimeter cells are made up of long (L) and short fibres (S) (see Section 3.2.4.1). The long fibres are sensitive to both electromagnetic (photons and electrons) and hadronic showers, while short fibres miss most of the electromagnetic (EM) showers. This is because the EM showers develop and dissipate most of their energy in the front part of the absorber, making S-fibres more sensitive to hadronic showers.

In the CMSSW release used for this analysis, the HF detector simulation does not accurately describe the ratio of the long to short fibre response. The deficiency in the description is due to imperfections in the modelling of the short-fibre response for low energy particles. Since different event generators (and tunes) are characterised by different energy spectra and particle compositions, the inaccurate modelling of the S-fibre response could influence the observed level of agreement among the different MC predictions. For example, if the primary difference in particle composition between the

tunes is attributed to low-energy π_0, the non-linearity effects of the calorimeter would be exaggerated by the use of $L + S$ fibres; this would lead to a larger disagreement amongst the different predictions. Accordingly, the analysis was repeated using L-fibres exclusively, i.e., measuring the energy of RecHits produced in the long fibres.

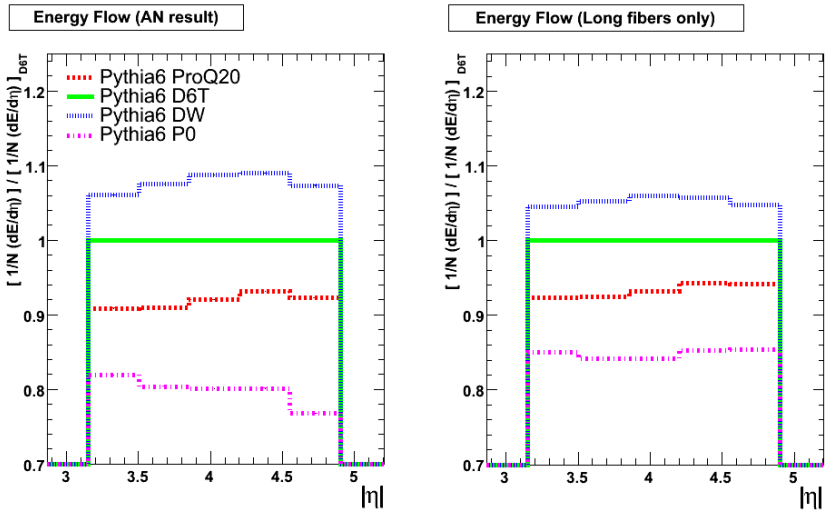

Figure 5.4: Energy flow using $L+S$ fibres and L-fibres - Ratio of energy flow predictions from various detector-level simulations for different tunes, with respect to PYTHIA6.4 D6T predictions, for long + short fibres (left) and long fibres only (right) [172].

Figure 5.4 shows the ratio of different MC simulations for minimum bias events at $\sqrt{s}$ = 900 GeV using both long and short fibres (left) and long fibres only (right). Figure 5.5 display the ratio of the two histograms for minimum bias events (left) and events with a dijet system (right), indicating that the relative agreement of different MC predictions may change by approximately 5% (10% in the last bin for tune P0) if only the long fibres of HF are taken into account. Figure 5.6 is the same as Figure 5.5, shown for $\sqrt{s} = 7$ TeV. The relative agreement would change by a smaller amount, for the higher centre-of-mass energy (3% - 5%).

The energy calibration of the HF ($L + S$) was done with single particle test beam data. The calibration imposed the requirement that the mean $L+S$ signal be 100 GeV for incoming electrons (or pions) of the same energy. The default HF calibration would correspond to $\sim$ 77 GeV ($\sim$ 60 GeV) for 100 GeV electrons (pions), if only L-fibre signals were used from the same test data. Moreover, the non-linearity corrections for

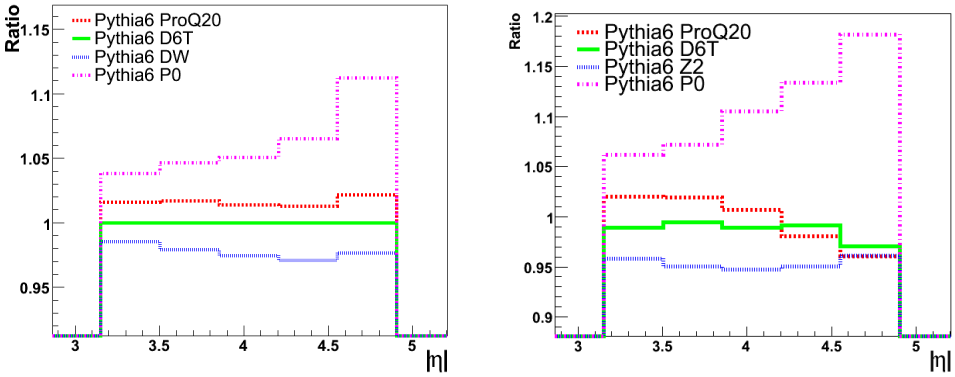

Figure 5.5: Ratio of energy flow using L+S fibres over L-fibres, $\sqrt{s} = 900$ GeV - Ratio between $L + S$-fibre and L-fibre ratios, for minimum bias events (left) and dijet events (right), $\sqrt{s} = 900$ GeV. [171] [172]

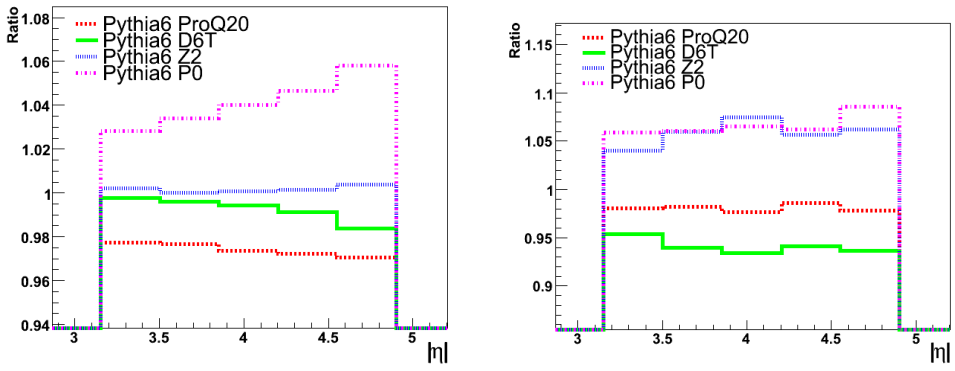

Figure 5.6: Ratio of energy flow using L+S fibres over L=fibres, $\sqrt{s} = 7$ TeV - Ratio between $L + S$-fibre and L-fibre ratios, for minimum bias events (left) and di-jet events (right), $\sqrt{s} = 7$ TeV. [171] [172]

only L-fibres would differ from the current ones. As a result, the interpretation of the energy flow measurement using long fibres exclusively would be difficult.

Furthermore, we know that low energy photons from π_0-decays produce signals primarily in the long fibres due to their decay channel, $\pi_0 \rightarrow \gamma\gamma$. Therefore, the precision in the simulation of the short fibres has a low impact on the measurement of the average energy flow.

With these considerations in mind, the energy flow was measured with the default HF setup. To cover the effect of the S-fibre imperfections, a systematic uncertainty is assigned to the measurement (see Chapter 7.2.8).

5.2.1.2 Non-uniformity Effects

The geometrical non-uniformity of the HF calorimeter is inaccurately modelled in the detector simulation. In the current description, the HF's active volume is described as a cylinder with a hole around the beam pipe without any non-sensitive or "dead" zones. The actual detector has a fraction of dead zones due to the mechanical structure of the detector's 20°–wedge assembly. The calibration constants are also affected, as the test beam was done in a way such that the electron or hadron shower was completely contained within the sensitive part of the wedge. In short, 100% occupancy is implemented in the simulations [176].

Until the HF geometry is updated and calibration constants are adjusted in the detector simulation, the effect of dead material in the HF rings must be accounted for in data. Implemented on a tower-by-tower basis, the values of the correction factors range from 0.98, for the innermost HF rapidity segment, to 0.90, for the outer most HF rapidity segment (Table: 5.3). They are applied to a tower's energy according to Eq. 5.2, with $i\eta$ defined in Chapter 3.2.4.1 [177].

$$Corrected\,Energy\,Flow(|i\eta|) = \frac{Energy\,Flow(|i\eta|)}{ratio(|i\eta|)}. \tag{5.2}$$

Table 5.3: Ratio of the active HF tower area to the nominal[131][145] tower area [177].

$\|i\eta\|$	30	31	32	33	34	35	36	37	38	39	40
Ratio	0.982	0.978	0.974	0.969	0.963	0.956	0.948	0.938	0.926	0.912	0.895

5.2.2 Treatment of Pile-up

We consider two types of pile-up events. The first is the trivial case where hard events produce multiple vertices, so pile up can be easily identified by looking at the number of reconstructed vertices. The second is pile-up events which consist of single diffractive dissociation (SD) and double diffractive dissociation (DD) events, which leave no tracks in the central detector but produce energy deposits in the forward detectors.

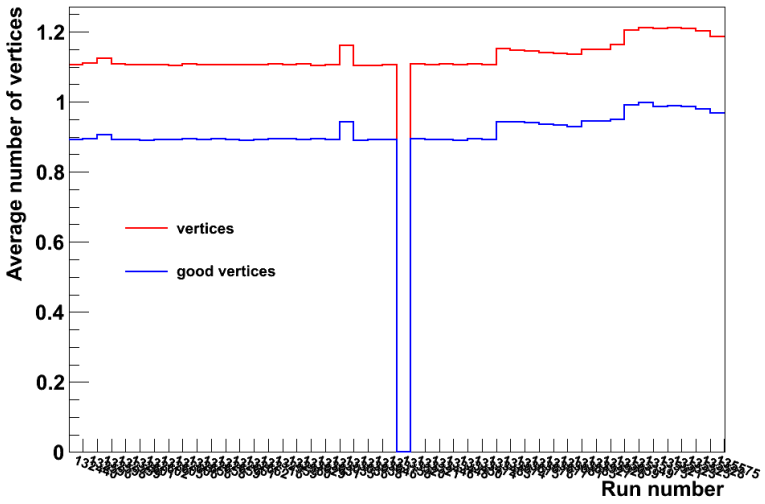

Figure 5.7: Number of vertices per event for each run - The red line shows total number of vertices in the event while the blue line shows the number of vertices that pass our vertex selection criteria. [171] [178]

Figure 5.7 shows the average number of vertices in an event as a function of run number, with the red line denoting the average of all vertices and the blue one representing all vertices that pass our vertex selection criteria (labelled as "good"). The first two bumps in Figure 5.7 are statistical fluctuations, occurring in runs where there are few events (150 events and 650 events, respectively). The dip occurs for a run with no events. After run 133483, we see a systematic increase in the number of vertices.

Runs are divided into segments called lumi-sections, in order to identify a particular run period. Each segment is 23.3 seconds long. The expected number of pileup events can be calculated using the instantaneous luminosity of each bunch crossing for each luminosity section, minimum bias cross-section (71.3 mb), and the circulation rate using the following formula (Eq. 5.3):

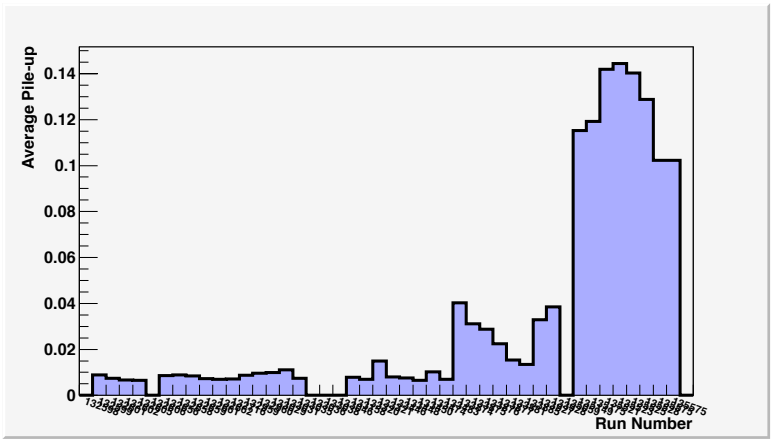

Figure 5.8: Average pile-up per event for each run - Calculated using instantaneous luminosity of each bunch crossing for each luminosity section (23.3 seconds), minimum bias cross-section (71.3 mb), and the circulation rate. [171] [178]

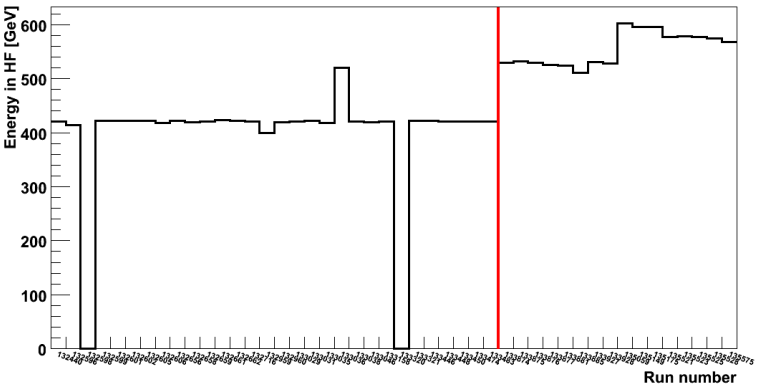

Figure 5.9: Total energy per event in the HF (both sides) for each run - The red line marks the last run used. Runs beyond the red line were discarded for this analysis. [171] [172]

$$N_{pileup\,_{xing,\,ls}} = \frac{L_{\,xing,\,ls} \cdot \sigma_{minimum\,bias}}{circulation\,rate}. \tag{5.3}$$

Figure 5.8 plots the result of this calculation for each run, with 1% of pile up events for runs we consider, and 2%–14% for the remaining runs in the data sample. The regions without pile-up events are runs where the database did not contain information on the luminosity, making the calculation impossible.

The total energy deposition in the HF stays relatively constant until run 133483 (Figure 5.9). The four 'kinks' seen can be attributed to statistical fluctuations and do not affect our measurement of the average energy flow in the HF. The increase in energy as of run 133483 can be interpreted as an increase in both hard and soft pile-up. In order to keep the effect of pile-up events to a minimum, we discard all runs beyond 133483 for the measurement of the HF energy flow.

5.2.3 Energy Flow

The energy flow is relatively insensitive to the polar angle between the interaction point and the detector, compared to the transverse energy flow (see Eq 5.4). Therefore, a measurement of the transverse energy flow would contain a larger systematic uncertainty, due to the possible misestimation of the exact interaction point and its distance from the detector. The energy flow's relative insensitivity to a displacement of the vertex (see Chapter 7) makes it the variable of choice for this investigation.

$$E_T = E \sin\theta \tag{5.4}$$

The energy is measured with the HF calorimeters, by summing all energy deposits in the HF towers above a threshold of 4 GeV. This threshold, determined previously for diffractive studies [179], suppresses the electronic noise; for the purpose of this study, the value is cross-checked using non-collision events for the purpose of this study. In addition, events in which particles hit the photo-multipliers and cause large signals in the HF calorimeter towers were removed from the analysis using dedicated algorithms. The rejection criteria are based on the topology of energy deposits and the pulse shape/timing of the signals in HF [180]. The resulting energy flow data are shown in five bins, grouping two consecutive rings together (see Chapter 3.2.4.1), with the $+\eta$ and $-\eta$ measurements averaged into a single distribution in $|\eta|$.

5.2.3.1 Energy Flow in Minimum Bias Events

Figure 5.10 shows the energy flow in the HF acceptance range for minimum bias events at two different centre-of-mass energies, $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV. The results for $\sqrt{s} = 0.9$ GeV are compared with several different model predictions. Different tunes for the parameters of the simulation of the underlying event in PYTHIA 6.4 [76] are used: D6T [66], DW [66]PRO-Q20 [95], Z2 [100] and Perugia-0 (P0) [96], the latter using the new multiple interaction model implemented in PYTHIA [103]. The generated events are processed through a full simulation of the CMS detector response. The results for 7 TeV are also compared with several MC predictions, coming from PYTHIA 8 [104], PHOJET [86], and four PYTHIA6.4 tunes: D6T, PRO-Q20, Perugia-0 (P0).

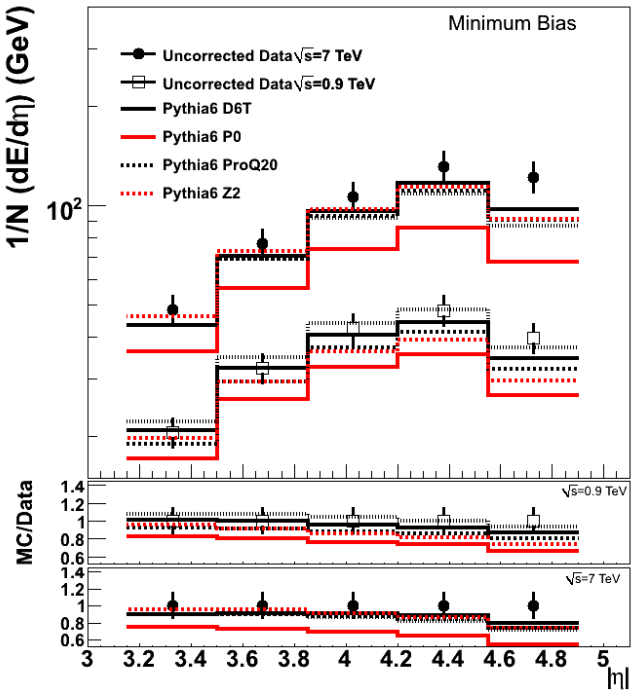

Figure 5.10: Uncorrected energy flow in the minimum bias sample - Energy flow as a function of $|\eta|$ at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV. Uncorrected data are shown as points. The histograms are predictions from different MC models. [171] [181]

The simulated samples were generated with varying number of events. The difference in the number of events generated leads to different statistical uncertainties for the MC distributions; they are not shown in the figures as they are negligible, given the size of the data sets. The systematic uncertainties are marked with vertical error

bars, are estimated to be 10% of the measured value in a given bin. A discussion of the systematic studies is presented in Chapter 7. The general trend is that the simulations predict less energy flow for both centre-of-mass energies but they predict the observed increase with respect to the increase in $\sqrt{s}$. A detailed discussion of the corrected data compared to various MC event generators can be found in Chapter 8.

5.2.3.2 Energy Flow in Di-jet Events

Figure 5.11 shows the energy flow in the HF acceptance range for events with a central di-jet system, at two different centre-of-mass energies, $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV. Along with the event generators and tunes used for comparisons in the minimum bias scenario, distributions from PHOJET[86] have been included. The statistical error for the di-jet data sets is not shown, as it is considered to be negligible.

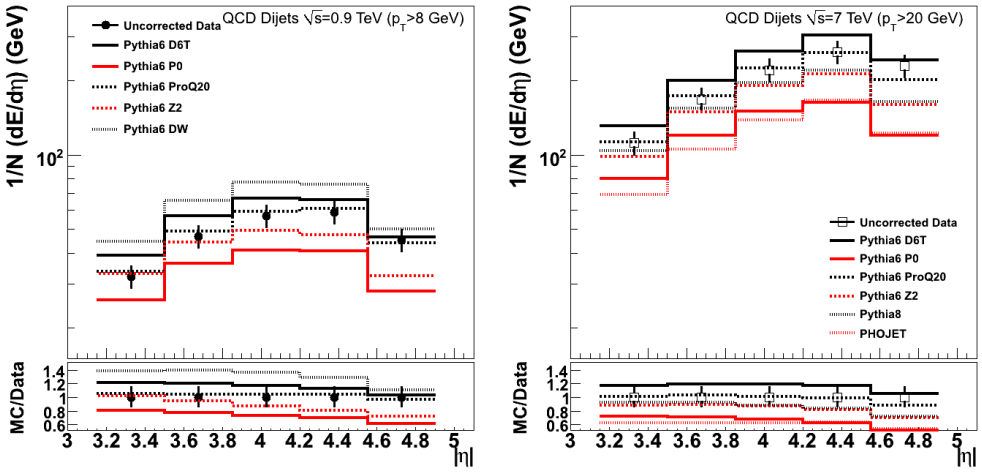

Figure 5.11: Detector-level energy flow, di-jet sample, $\sqrt{s} = 900$ GeV - Energy flow in the di-jet sample, as a function of $|\eta|$ at $\sqrt{s} = 0.9$ GeV (left) and $\sqrt{s} = 7$ GeV (right). [171] [181]

Here, the predictions do a better job than for minimum bias events, for $\sqrt{s} = 7$ TeV, with one distribution (PRO-Q20) describing the data in most bins. For the 900 GeV samples, the predictions encompass the data but not one tune describes the distribution. A comprehensive discussion on the comparison between data and simulations can be found in Chapter 8.

6

Correction to Hadron Level

The purpose of an unfolding procedure is to correct the detector-level measurements back to the hadron level, by removing artificial effects caused by the apparatus and trigger. Bin-by-bin unfolding [182] is the most commonly used procedure at most experiments. It works by scaling the i'th bin of the detector-level measurement by a factor calculated from a correlation matrix of the detector response. Often a full unfolding is preferred instead, allowing for a more detailed and complete response matrix. However, this measurement is not a counting measurement and there were no directly comparable hadron-detector level objects, rendering such an approach unfeasible.

Here, data are corrected to hadron level using multiple tunes of the event generator PYTHIA 6.4, together with a simulation of the CMS detector based on GEANT4. The position and width of the beam spot in the simulation are adjusted to that determined from the data. The detector simulated events pass through the same analysis chain as data. Corrections are applied on a bin-by-bin basis, correcting for inefficiencies, acceptance and bin-to-bin migration arising from the finite detector resolution. The correction factors are calculated as the ratio of the event simulation predictions on hadron and detector level.

6.1 Hadron-level Selection

Hadron level is defined by the presence of stable particles (unstable: $\tau < 10^{-12}$ s) and the exclusion of neutrinos and muons. Neutrinos are excluded as their energy cannot be measured by the calorimeter. While muons do leave a minimum ionising signal

(mip) in the calorimeter, their contributions would be too small to the overall energy flow measured with a calorimeter. Figure 6.1 shows that the multiplicity and energy is low in this region. The most important point is the fractional energy of the muons - Figure 6.2 shows that they would not account for a large fraction of the energy in the HF. This confirms that muon contribution to the energy flow (on hadron level) is negligible, allowing us to exclude it from the correction factor.

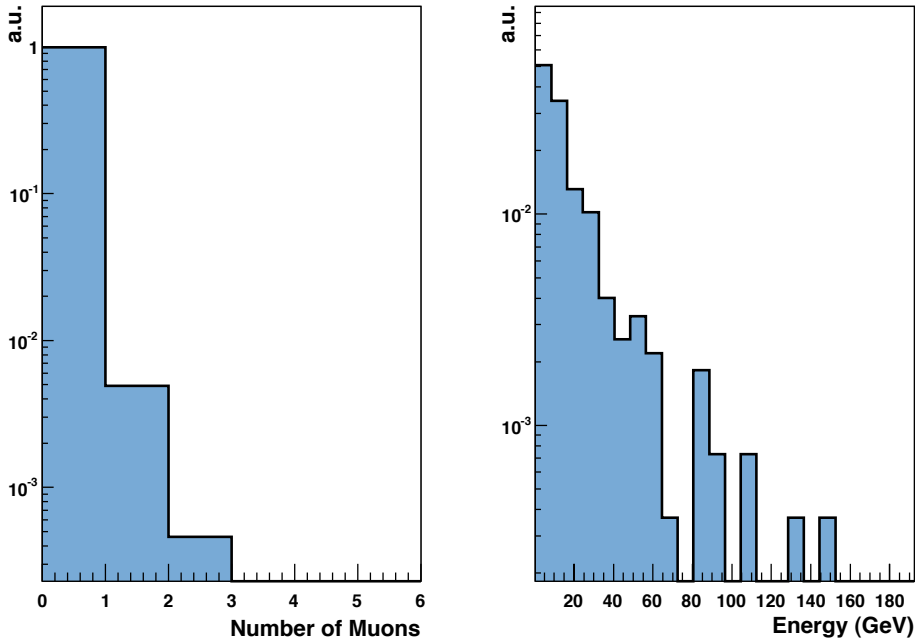

Figure 6.1: Multiplicity and energy spectrum of hadron-level muons within the measurement region - Number of muons (left) and energy spectrum (right) within the measurement region, $3.152 < |\eta| < 4.902$) on hadron level, normalised to the integral of the histogram, $\sqrt{s} = 7$ TeV. The events were generated using PYTHIA 6.4 Tune D6T.

Furthermore, we require the presence of a charged particle within the BSC range ($3.9 < |\eta| < 4.4$) on both sides, to emulate the BSC1 trigger used on the detector level ($3.9 < |\eta| < 4.4$) [166]. That is, we include everything except single diffractive dissociation and elastic processes for the correction. No other bias on the particle energy or rapidity is introduced.

In the case of the dijet systems, we use hadron-level jets defined by the anti-k$_T$ algorithm ($R = 0.5$) with identical selection criteria applied on the kinematics (p_T, η,

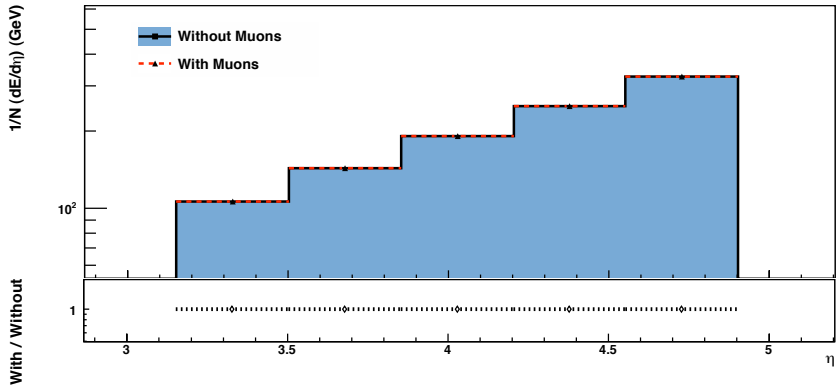

Figure 6.2: Muon contribution to the energy flow, on hadron level - Energy flow in the HF (both sides) with and without muons, on hadron level, $\sqrt{s} = 7$ TeV. The events were generated using PYTHIA 6.4 Tune D6T. The contribution from muons is negligible, as illustrated by the figure and the ratio of the two distributions.

ϕ) to both detector-level and hadron-level jets.

6.1.1 Impact of Hadron-level Selection

The kinematic region of the measurement as defined at the hadron-level is highly correlated to that defined at the detector-level. In order to minimise the sensitivity of the measurement to model dependence, the kinematic regions for both the measured and simulated events should be similar. If this is not the case, the final measurements will be the result of extrapolations based on Monte Carlo predictions. In this analysis, the reconstructed quantities used in the selection process exhibit little bias after corrections. The hadron-level kinematic region is, therefore, very similar to that at the detector-level. However, the minimum bias triggers may introduce some inefficiencies.

In order to minimise the possible bias and ineffeciencies that arise due to the BSC-trigger, different hadron-level selections are explored. The ideal scenario would be where the hadron-level selection that most resembles the actual trigger, i.e., one charged particle within the BSC acceptance range, produces the exact same correction factor as DET-LEVEL BSC TRIGGER. Such agreement would indicate 100% efficiency of the trigger and emulator.

6.1.1.1 Impact of hadron-level selection on correction factors

Figure 6.3–6.4 show the correction factors for different hadron-level selections, comparing them to the correction factor obtained with the detector-level selection. The conditions were as follows:

- DET-LEVEL BSC TRIGGER: the hadron-level distributions were obtained by applying the standard detector-level triggers on the generated sample;
- HAD LEVEL: selection criteria applied on hadron-level objects (generated particles), to simulate detector-like requirements, i.e, a particle within a given η-range;
- MEASUREMENT ACCEPTANCE: generated particles within the measurement acceptance (on each side), with total energy greater than zero;
- BSC ACCEPTANCE: similar to MEASUREMENT ACCEPTANCE, but using the reduced η-range of the BSC;
- CHRG PART: replaces the energy requirement with the necessity of a charged particle within the specified η–range. The BSC-trigger requires one charged particle.

The plot for the minimum bias scenario (Figure 6.3) demonstrates that the requirement of some energy deposit within the BSC η-range on hadron level gives the correction factor closest to that obtained using a detector-level trigger. Requiring activity within the measurement acceptance produces a smaller correction factor, with no difference between an energy requirement or a particle condition. Imposing a condition on the particle instead of energy within the BSC acceptance gives us a larger correction factor, with no difference observed due to changes in energy threshold.

The dijet scenario plot (Figure 6.4) indicates that the different hadron-level selection criteria do not affect the correction factors as much as in the minimum bias scenario. In fact, with the exception of the BSC-like (charged particle, $3.9 < |\eta| < 4.4$) all of them produce the same result. This is due to the hard scale defined by the presence of a dijet in the event, which increases the density of particles in the forward region; the increased particle multiplicity makes it equally probably to find a particle for each of the limitations. That is, finding a charged particle is as probable as finding a particle $E_{particle} > 0.2$ GeV.

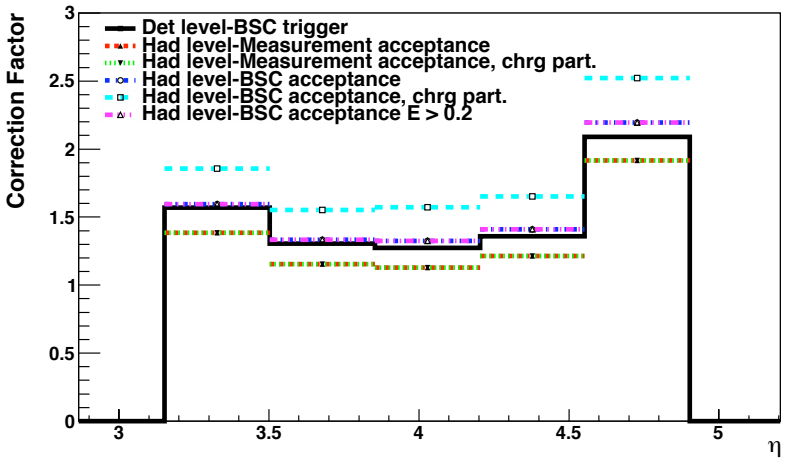

Figure 6.3: Correction factor for minimum bias events, $\sqrt{s} = 7\ TeV$ - Correction factors for the different hadron-level selection criteria compared to the correction factor obtained from detector-level selection criteria for minimum bias events for $\sqrt{s} = 7\ TeV$, using PYTHIA6.4 D6T. The "measurement acceptance" histograms overlap with each other, as do most of the "bsc acceptance". [171]

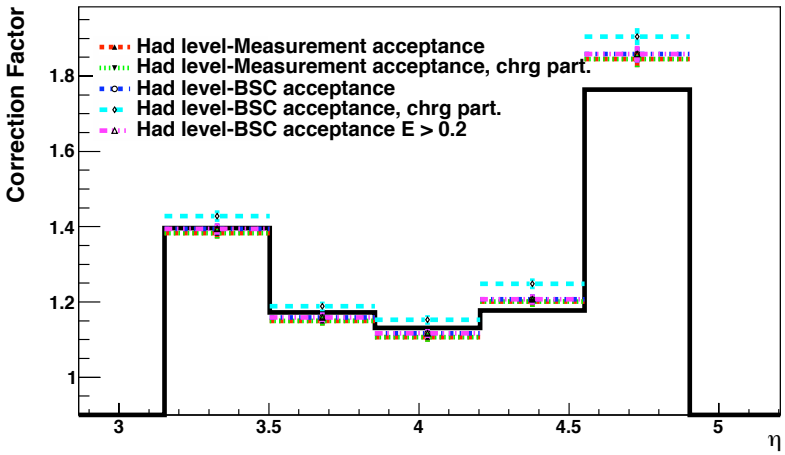

Figure 6.4: Correction factor for dijet events, $\sqrt{s} = 7$ TeV - Correction factors for the different hadron-level selection criteria compared to the correction factor obtained from detector-level selection criteria for minimum bias events for $\sqrt{s} = 7$ TeV, using PYTHIA6.4 D6T. All the lines (except the BSC-charged particle) overlap. [171]

Looking at the single diffractive component that survives the hadron-level selection criteria (kinematic restrictions) in minimum bias events (Table 6.1), we find that twice the single diffractive dissociative (SD) component remains, when using the measurement acceptance for the trigger instead of the BSC limits. That is, the stricter, BSC-like requirement rejects twice as many SD events as the full measurement range (measurement HF or MHF). By allowing a larger fraction of diffractive events, the MHF selection results in smaller overall correction factors than both the actual detector-level trigger and the BSC-like requirements. That is, the increase in the fraction of diffractive events causes the difference in the correction factor between the various hadron level definitions, an effect which disappears in the dijet scenario as there is a negligible rate of diffractive events ($\sim 1\%$).

Table 6.1: Fraction of single diffractive events that survive the hadron-level "trigger" with a charged particle condition, for minimum bias events, $\sqrt{s} = 7$ TeV.

Tune	Measurement-acceptance	BSC-acceptance
D6T	7%	3.3%
PROQ20	7%	3.3%
P0	6.9%	3.4%
Z2	7%	3.4%

6.1.1.2 Impact of hadron-level selection on model dependence

A correction factor calculated using Monte Carlo simulations carries an inherent dependence on the model used (model dependence). The possible bias due to the correction factor's reliance on the underlying description of physics can be estimated by comparing it to one (e.g. a default standard) or multiple models and gauging the deviation. Figures (6.5–6.6) show the variation of the model dependence of the correction factors (D6T used as reference, shown in black) for the various hadron-level selections. The model uncertainty is smallest for the 7 TeV minimum bias scenario using the charged particle criteria within the BSC-range ($\approx 4\%$), almost half the observed uncertainty when using the full measurement range ($\approx 7\%$). For dijet events, the model uncertainty increases to 15% for the BSC-like "trigger" and $\approx 10\%$ for the measurement acceptance.

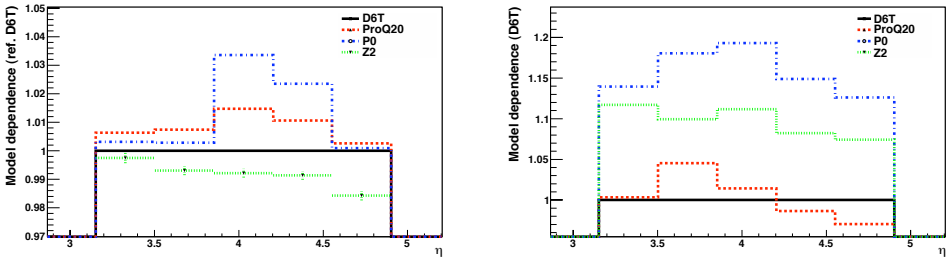

Figure 6.5: Model dependence with BSC-selection for minimum bias and dijet events, $\sqrt{s} = 7$ TeV - Model dependence of the correction factors, w.r.t. the correction factor obtained using PYTHIA6.4 D6T, for BSC-like hadron-level selection criteria, for minimum bias (left) and dijet (right) events. Note the different scales for the two event types. [171]

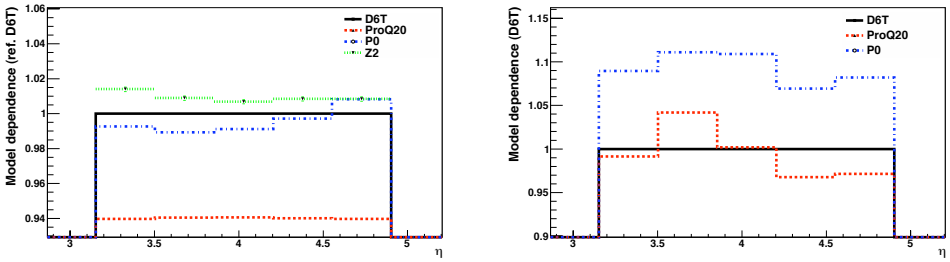

Figure 6.6: Model dependence with measurement-acceptance selection, for minimum bias events and dijet events, $\sqrt{s} = 7$ TeV - Model dependence of the correction factors, w.r.t. the correction factor obtained using PYTHIA D6T, for hadron-level selection criteria using the measurement acceptance instead of the BSC acceptance, for minimum bias (left) and dijet (right) events. Note the different scales for the two event types. [171]

PYTHIA 6.4 MPI tunes do not differ much in their treatment of soft diffraction, and, regardless of the selection criteria, the various parameter sets predict comparable fractions of diffractive events (Table 6.1). Therefore, assuming that neither the fraction nor the modelling of the diffractive component is the reason for the increased model dependence in the MHF samples seems justified.

Seeing that the HF or measurement acceptance requirement is the least restrictive, it is reasonable to assume the correction factor accounts better for an extrapolation to a larger event sample than in the other cases. In particular, the extrapolation depends on the additional particles for which it corrects: particles outside the HF acceptance and neutral particles. Therefore, the model dependence of these correction factors increases, as models differ in their description of particle composition and spectra. However, this is not the case in the dijet selection. The presence of a dijet system in the central region supersedes the kinematic limits in the forward region; the jets set the hard scale, and thus drive the particle and energy flow.

The choice to minimise the model dependence for $\sqrt{s} = 7$ TeV minimum bias samples was the deciding factor in setting the hadron-level criteria.

6.2 Correction Factors

In the first four bins, the correction factors are around 1.60 – 2.50 (1.50 – 2.50) for the $\sqrt{s} = 900$ GeV ($\sqrt{s} = 7$ TeV) minimum bias data. Due to the presence of dead material (beam-pipe) at $|\eta| \approx 4.7$–4.9, which absorbs soft particles, the correction factor for the highest rapidity bin is larger (≈ 2.5). The overall lower correction factors for the $\sqrt{s} = 7$ TeV data can be related to the fact that higher energy particles, which are produced at higher centre-of-mass energies, are to a less degree affected by additional dead material in front of the detector.

These are two separate issues. The last bin has the highest correction factor for all scenarios because of the additional beam-pipe / dead material. But the differences in the overall correction factor has an energy scale dependence; 900 GeV Min Bias $>$ 900 GeV dijets $>$ 7 TeV Min Bias $>$ 7 TeV dijets. This indicates that the detector response improves for higher energies, leading to a smaller correction factor as the energy scale increases.

The correction factors from the bin-by-bin method (Figures 6.7–6.8) are model dependent. Moreover, they are affected strongly by migrations. For example, if a model is used in which there are large migrations (e.g. steeply falling p_T spectrum), one would get a correction factor very different from one that was obtained using a model in which 'little' is outside a bin. As a result, this method of correction can only be applied if the simulation model describes the reconstructed distributions well and the migration effects are minimal.

The correction factors used for correcting the data are obtained by taking an average of the four distributions shown in Figures 6.7–6.8.

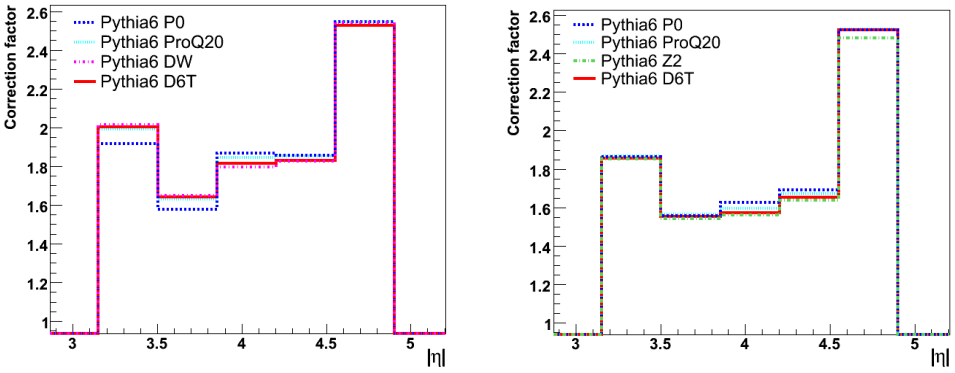

Figure 6.7: Correction factor for minimum bias events, $\sqrt{s}$ = 900 GeV and $\sqrt{s}$ = 7 TeV - The final correction factors are derived from an average of the different PYTHIA 6.4 tunes shown here. Left: $\sqrt{s} = 900$ GeV. Right: $\sqrt{s} = 7$ TeV. [171] [172]

6.3 Migrations

Migration effects between different pseudorapidity bins may influence the energy flow measurement and have to be taken into account for the corrections at hadron level. It is possible, for example, that particles (e.g. low p_T charged particles) generated outside the acceptance of HF deposit energy in the detector due to scattering in regions of dead material or due to bending caused by the magnetic field.

We see that for the average energy flow, we have reasonable correction factors (Figures 6.7 – 6.8), but distributions of the energy spectra (Figure 6.9) do not look very similar when comparing simulations on hadron-level to detector-level simulations.

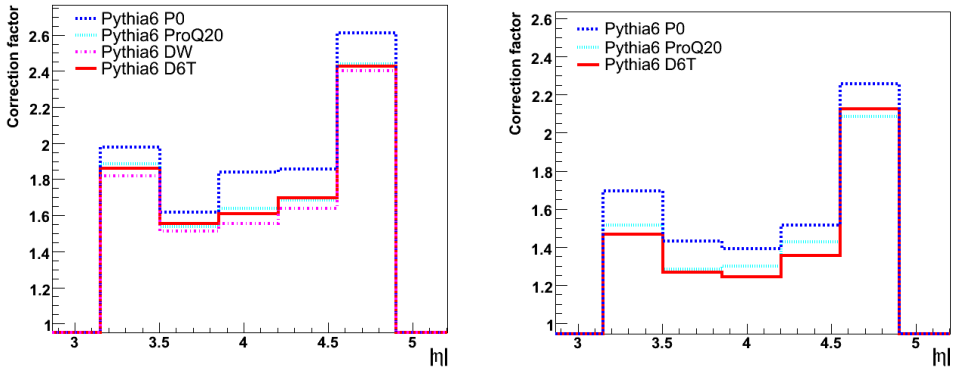

Figure 6.8: Correction factor for dijet events, $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV - The final correction factors are derived from an average of the different PYTHIA 6.4 tunes shown here. Left: $\sqrt{s} = 900$ GeV. Right: $\sqrt{s} = 7$ TeV. [171] [172]

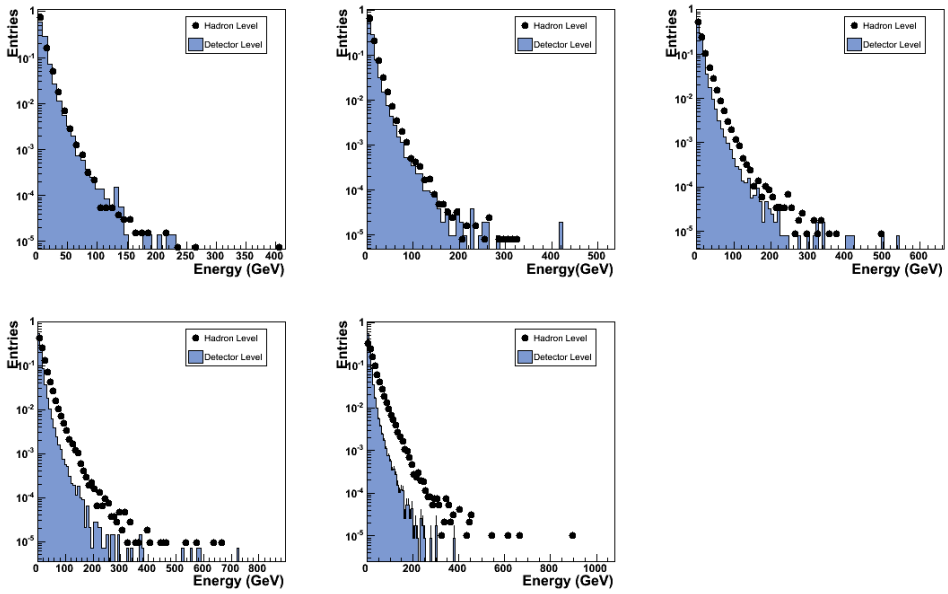

Figure 6.9: Energy spectra on hadron level and detector level, $\sqrt{s} = 7$ TeV, di-jet samples - HF energy spectra in bins of η used for the measurement of the energy flow. Plots compare hadron-level to detector-level simulations, $\sqrt{s} = 7$TeV, for a di-jet system. Each plot corresponds to one bin of the measurement shown in the result plots, i.e., first bin ($3.152 < \eta < 3.503$) is top left to the last bin ($4.730 < \eta < 4.903$) at bottom-right. [171] [172]

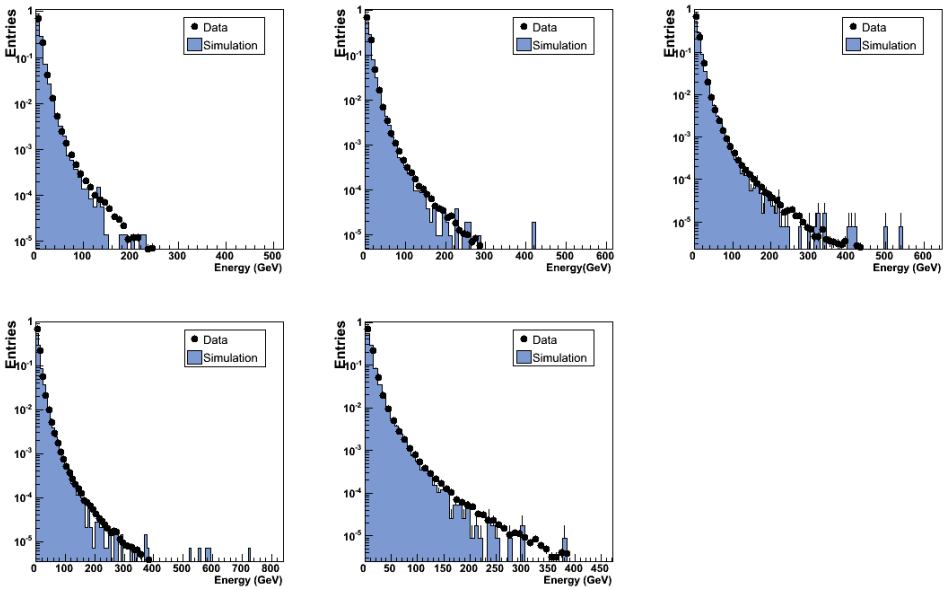

Figure 6.10: Energy spectra from data compared to detector-level simulated events, $\sqrt{s}$ = 7 TeV, di-jet samples - HF energy spectra in bins of η used for the measurement of the energy flow. Plots compare detector-level simulations to data, $\sqrt{s}$ = 7TeV, for a di-jet system. Each plot corresponds to one bin of the measurement shown in the result plots, i.e., first bin ($3.152 < \eta < 3.503$) is top left to the last bin ($4.730 < \eta < 4.903$) at bottom-right. [171] [172]

Uncorrected data, on the other hand, agrees well with the detector-level simulations (see 6.10). The discrepancies suggest that the differences are due to detector-related effects, such as:

- energy resolution and non linear response in the HF;
- absorption of particles in dead material in front of the HF;
- bending of particles from the central region into the HF region due to magnetic field, and bending particles out of the HF region due to the magnetic field.

The energy resolution and non-linear response is simulated in the detector MC and was cross-checked against test beam data. Other than varying the material budget (done previously [183]), there is no way to check the absorption of particles by the dead material in front of the HF. The study showed that the effects were negligible, except for the last bin of the measurement as it is closest to the beam pipe (discussed further in 6.3.2 and 7.1.5). This dead material (beam pipe) accounts for the clear difference between the hadron level and detector level simulation seen primarily in the last bin of Figure 6.9,

In order to account for the magnetic field effects on soft particles throughout the detector and measurement range, we try to identify the fraction of particles and energy that enters the HF, compared to what an event generator predicts on hadron level. This is covered in the following section.

6.3.1 Single Particle Response

RecHits were defined in 4.2.2 as the most basic reconstructed object of a calorimeter. By associating RecHits to the particle that fired them via a matching algorithm (see Figure 6.11), one can relate the HF response as a function of η, p_T and energy, to the properties of the generated particle; i.e. before the particle went through the detector simulation.

Here we use single particles; pions, electrons and photons generated using a "particle-gun" are studied separately, to obtain a better understanding of the HF's response to electromagnetic and hadronic particles.

Figure 6.12 shows the uniform distribution in pseudorapidity ($-6 < \eta < 6$) and transverse momentum (0.2 GeV $< p_T <$ 40 GeV) of the generated particles. The

Figure 6.11: Flowchart for matching RecHits to generated particles - The matching is done by mapping the reference numbers (ID) of each of the objects. The simulated tracks are matched to the generated particles by comparing the particle ID in the track to the particle ID of the generated particle.

extension of the phase space to $-6 < \eta < 6$ was done in order to include an overlap region with the CASTOR calorimeter at the $-z$ side, thereby allowing us to study back-scattering effects due to the presence of the calorimeter.

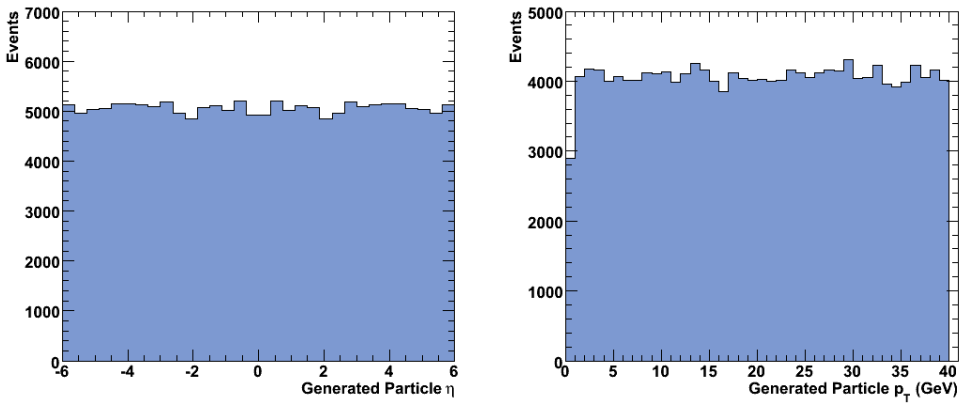

Figure 6.12: Pseudorapidity (η) and transverse momentum ((p_T) of generated particles using a single-particle gun - Left: Flat distribution, particles generated evenly for $-6 < |\eta| < 6$, using a single-particle gun. Right: Flat distribution, particles generated evenly for 0 GeV $< p_T <$ 40 GeV, using a single-particle gun. [171] [172]

6.3.1.1 Pions

The plot showing the pseudorapidity of the matched HF-RecHits versus the pseudorapidity of the generated particles (Figure 6.13) illustrates the fact that there are migrations (most clearly visible in the red region) in η. Moreover, one sees that there are pions beyond the HF's η–acceptance which enter the HF region and fire RecHits for both the high and low η bins. Lastly, we see a pronounced increase in migrated RecHits at $\eta \approx 4.9$, which is due to the presence of dead material (see Section 7.1.5).

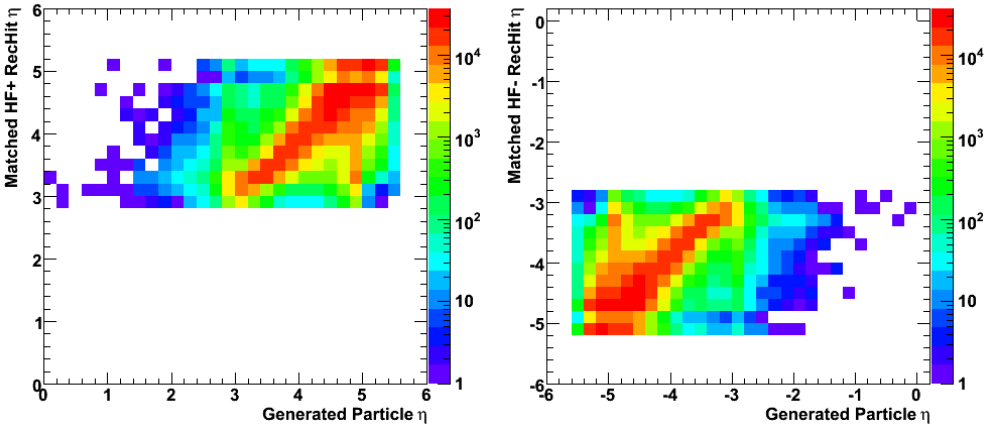

Figure 6.13: Pseudorapidity of the matched HF RecHits versus the pseudorapidity of the generated pions. - HF+ (left) and HF− (right). [171] [172]

In the left plot shown in Figure 6.14, the multiplicity of the HF RecHits is shown as a function of the pseudorapidity of the generated pion that produced the signal. In the right plot shown in Figure 6.14, the transverse momentum (p_T) of generated pions that produced the signal are shown as a function of the particle's pseudorapidity (η).

Three traits are instantly noticeable: the slant or slope in the left plot in Figure 6.14 shows that there are more RecHits with increasing pseudorapidity; the colours between $1 < |\eta| < 3$ in the right plot in Figure6.14 indicate that pions from outside the detector acceptance are equally distributed in p_T; for $3 < |\eta| < 4.5$, less low-p_T pions ($0 < p_T < 10$) from the front part of the calorimeter fire RecHits than particles with high transverse momentum (6.14: right).

6.3.1.2 Electrons

The plots for electrons (Figures 6.15 – 6.16) exhibit the same traits as seen in Figures 6.13–6.14, including the increase in RecHits near the last bin of the energy flow measurement ($\eta \approx 4.9$).

6.3.1.3 Photons

Figure 6.17 shows that there are comparably less RecHits due to photons originating outside the HF acceptance, than for electrons and pions. The dead-material effect ($\eta \approx 4.9$) seen in the case of electrons and pions remains prominent.

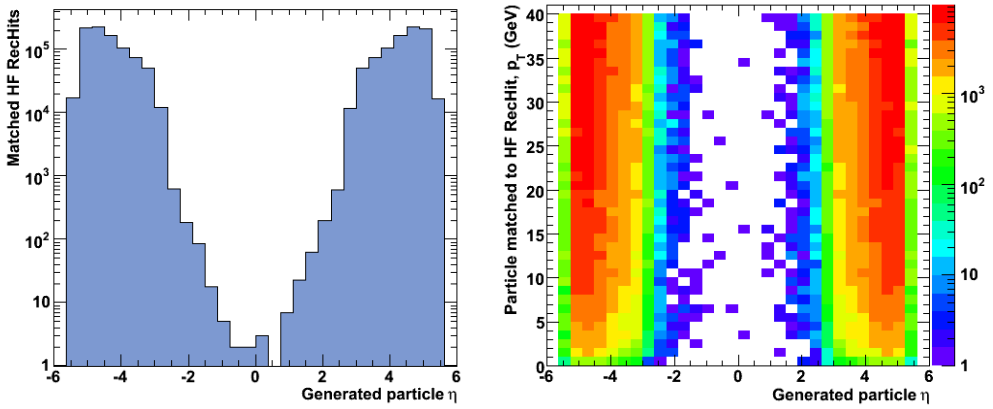

Figure 6.14: Multiplicity and kinematics of the HF RecHits matched to the generated pions that produced the signal - Left: RecHits matched to a generated pion versus the pseudorapidity of the generated particle. Right: Transverse momentum of the pions matched to HF RecHits versus the pseudorapidity of the generated particle.

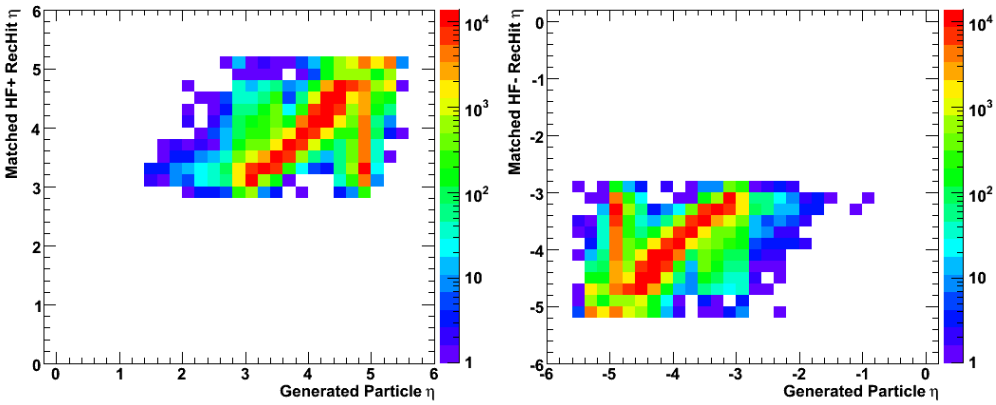

Figure 6.15: Pseudorapidity of the matched HF RecHits versus the pseudorapidity of the generated electrons. - HF+ (left) and HF− (right) [171] [172]

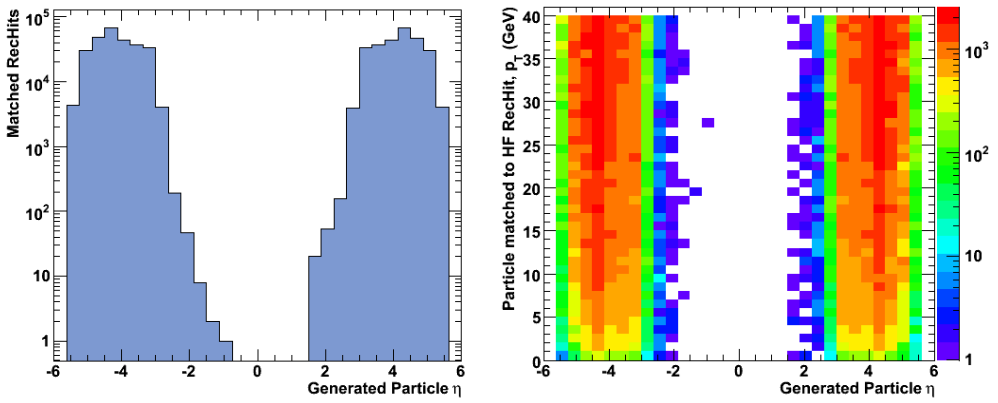

Figure 6.16: Multiplicity and kinematics of the HF RecHits matched to the generated electrons that produced the signal - Left: RecHits matched to a generated electron versus the pseudorapidity of the generated particle. Right: Transverse momentum of the electron matched to HF RecHits versus the pseudorapidity of the generated particle. [171] [172]

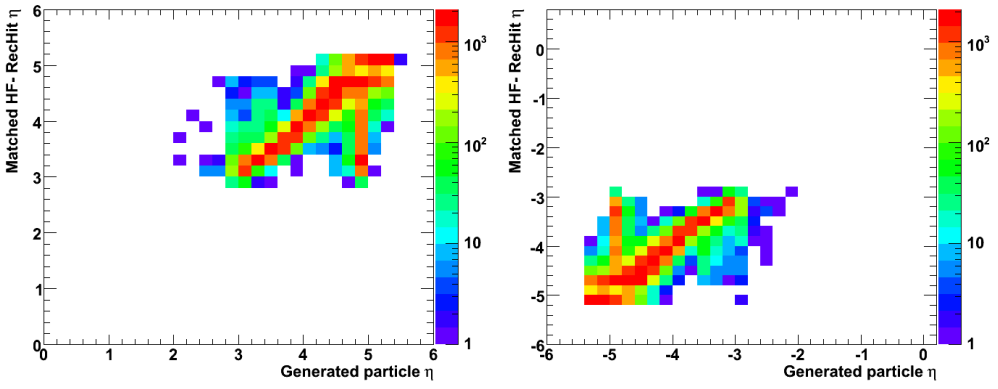

Figure 6.17: Pseudorapidity of the matched HF RecHits versus the pseudorapidity of the generated photons. - HF+ (left) and HF− (right). [171] [172]

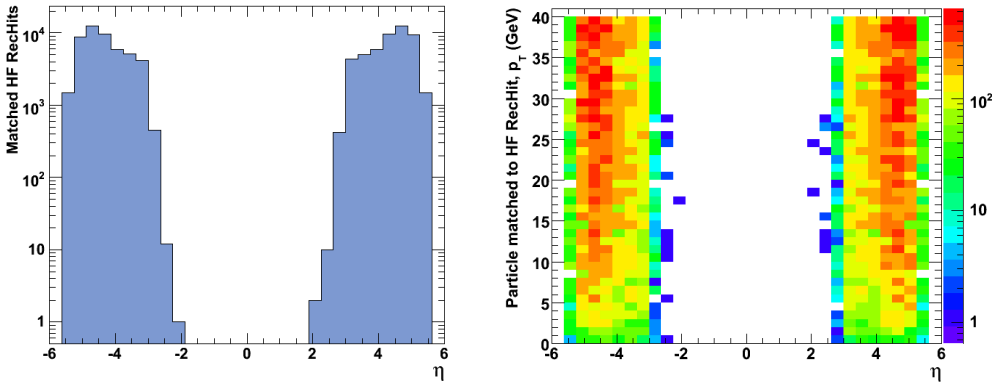

Figure 6.18: Multiplicity and kinematics of the HF RecHits matched to the generated photons that produced the signal - Left: RecHits matched to a generated photon versus the pseudorapidity of the generated particle. Right: Transverse momentum of the photon matched to HF RecHits versus the pseudorapidity of the generated particle. [171] [172]

Figure 6.18 (for photons) shows the same distributions as in Figure 6.14 (pions) and similar traits are noticeable. That is, an increase in RecHits in the more forward region of the calorimeter (slope in the left figure) or less low-p_T photons in the front part of the calorimeter. However, one difference is that there are little to no particles from beyond $|\eta| < 2$. This suggests that particles (pions and electrons) from beyond that pseudorapidity region arrive at the calorimeter due to the magnetic field while particles closer to the face of the calorimeter produce signals due to back-scattering and dead material effects.

Altogether, this suggests that the largest factor causing migrations is the presence of dead material and not the magnetic field.

6.3.2 Overall Effect of Migrations on the Energy Flow

Figure 6.19 shows two distributions for minimum bias events at $\sqrt{s} = 900$ GeV, using the same RecHit–Single Particle matching algorithm as before; the HF RecHit multiplicity versus the pseudorapidity of the generated particle that fired it, and the pseudorapidity of the matched HF RecHits versus the pseudorapidity of the generated particles. The main difference between the distributions for single-particle samples and the minimum bias samples is that there is a larger contribution to the HF RecHits from particles outside

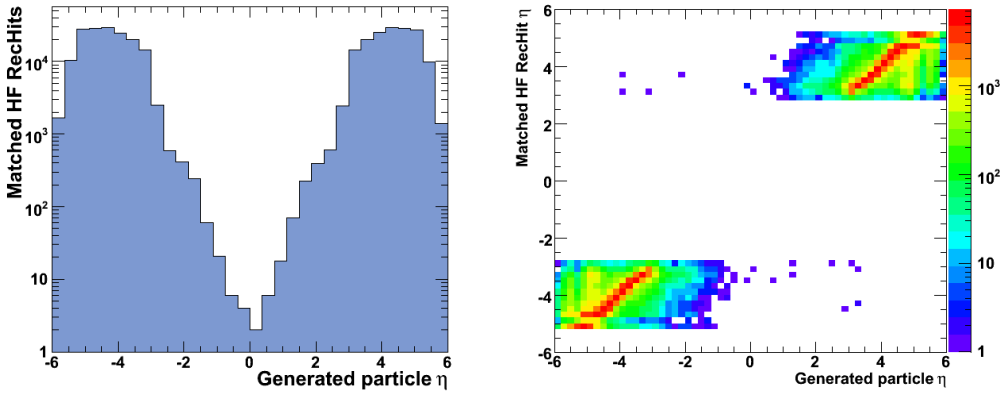

Figure 6.19: Multiplicity and kinematics of the HF RecHits matched to the generated particles that produced the signal in minimum bias events, $\sqrt{s} = 0.9\ TeV$ - Left: RecHits matched to a generated particle versus the pseudorapidity of the generated particle. Right: Transverse momentum of particles matched to a HF RecHits versus the pseudorapidity of the generated particle. [171]

the HF acceptance in minimum bias events than events simulated with a single-particle gun.

The contribution of particles from different areas outside the measurement acceptance to the total energy flow (HF+ and HF−) is compared in Figure 6.20. The plot shows the largest contribution ($\approx$ 20% for the last bin) to be from particles within ($4.9 < |\eta| < 5.2$), followed by particles from $\eta > 5.2$ ($\approx$ 10% for the last bin), and finally about 3% for the first and second last bins from particles within the HF η–range. Energy from particles in front of the calorimeter can be considered to be negligible.

Despite the large energy contribution from migrating particles to the last bin of our measurement, we do not assign an uncertainty due to this effect. The correction factor accounts for detector effects to a large extent, and we conclude that uncertainties attributed to the model dependence and dead-material effects would account for the remainder.

As there are no hadron-level objects that can be compared to detector-level objects in this measurement (e.g. particles cannot be compared to RecHits), it was not possible to perform investigations on the purity and stability of the correction factor. In the absence of an analysis, we use these studies to conclude that overall migrations are small enough to be negligible and that we can rely on the correction factors obtained.

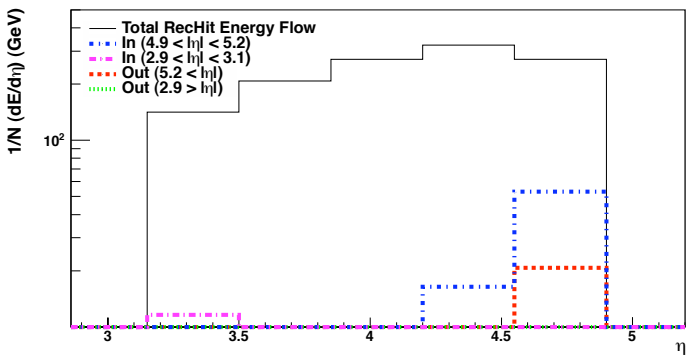

Figure 6.20: Effect of migrations on the overall energy flow, in minimum bias events, at $\sqrt{s} = 7$ TeV - Contribution to the energy flow from particles outside the measurement's acceptance (red, blue, green and purple) compared to the total energy flow (black), using RecHits, for minimum bias events at $\sqrt{s} = 7$ TeV. 'Out' signifies the particle was from outside the HF acceptance while 'In' means that it is from within the detector's range but outside the limit of the measurement.

7

Systematic Studies

Systematic studies can be split into two categories: systematic checks and systematic errors. The former consist of checks to test the internal consistency of a method and reveals potential mistakes. The latter involves evaluating the uncertainties in a measurement that arise from the continual misestimation of a quantity within the method. If a systematic check highlights a flaw, the method must be corrected or the uncertainty arising due to it must be evaluated. In the case of systematic uncertainty, estimate on the possible ambiguity is usually assigned to account for the persistent deviation.

7.1 Systematic checks

7.1.1 Consistency between previous and current measurements

Figures 7.1–7.2, compare results shown in [183] and the analysis done for this thesis. The deviations arise due to a variety of factors. New corrections (see Section 5.2.1.2) are implemented for each individual tower, which accounts for a difference of $1\% - 10\%$, depending upon the bin. Another 1% is accounted for by the change in automated algorithms and the reconstruction, covered individually (e.g. PMT hit removal algorithms, Section 7.2.5).

7.1.2 Splash Triggers

In addition to the standard minimum bias triggers used by CMS, splash triggers (triggers 42 and 43), requiring additional hits in a given part of the BSC (inner ring) were available. These triggers were used by the HCAL for calibration and timing studies.

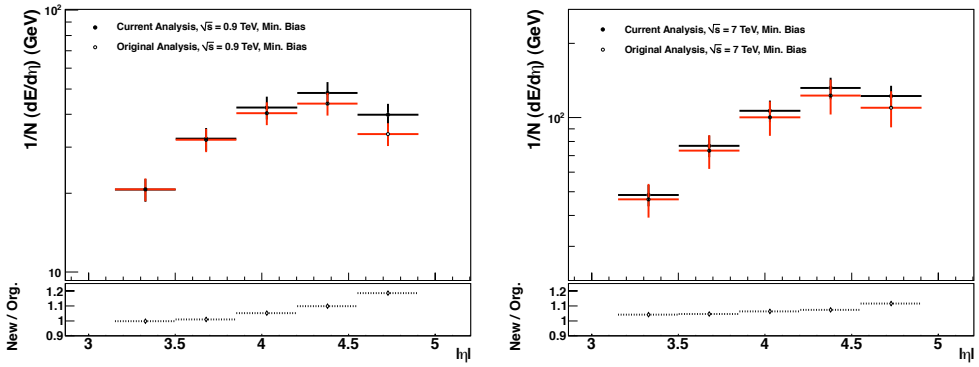

Figure 7.1: Comparison between previously released and current measurements in minimum bias events - Comparison of energy flow in minimum bias events, $\sqrt{s}$ = 0.9 TeV(left) and $\sqrt{s} = 7$ TeV (right), presented previously [183] using CMSSW356 (in red, open circles) and latest analysis done with CMSSW383 (in black, filled circles). The previous analysis for $\sqrt{s} = 0.9$ TeV is with runs taken in 2009 while the current analysis uses runs from 2010. Upto 10% of the difference in the last bins is due to the new tower corrections implemented. The uncertainties shown correspond to the energy scale uncertainty, 15% and 10% for the previous and current analyses, respectively. The data points differ from each other by approximately 1%–17% and 4%–10% at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, respectively. [171] [172]

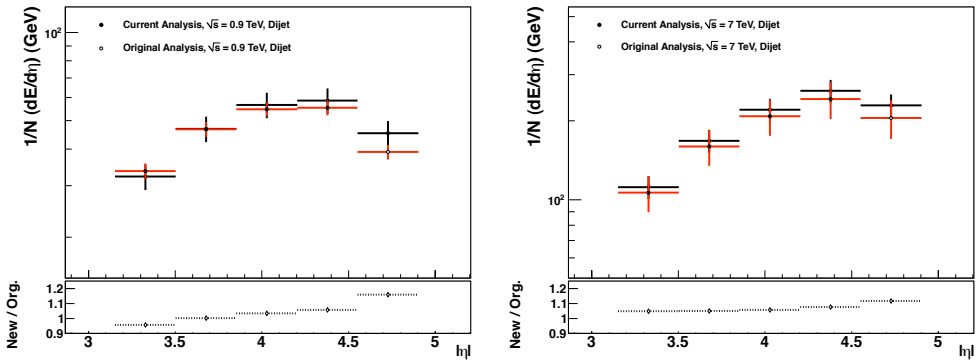

Figure 7.2: Comparison between previously released and current measurements in dijet events - Comparison of energy flow in dijet events, $\sqrt{s} = 0.9$ TeV(left) and $\sqrt{s} = 7$ TeV (right), presented previously [183] using CMSSW356 (in red, open circles) and latest analysis done with CMSSW383 (in black, filled circles). The previous analysis for $\sqrt{s} = 0.9$ TeV is with runs taken in 2009 while the current analysis uses runs from 2010. Upto 10% of the difference in the last bins is due to the new tower corrections implemented. The uncertainties shown correspond to the energy scale uncertainty, 15% and 10% for the previous and current analyses, respectively. The data points differ from each other by approximately 3%–15% and 4%–10% at $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, respectively. [171] [178]

Figure 7.3 shows that the application of this trigger has a negligible (less than 0.5%) effect on the measured energy flow.

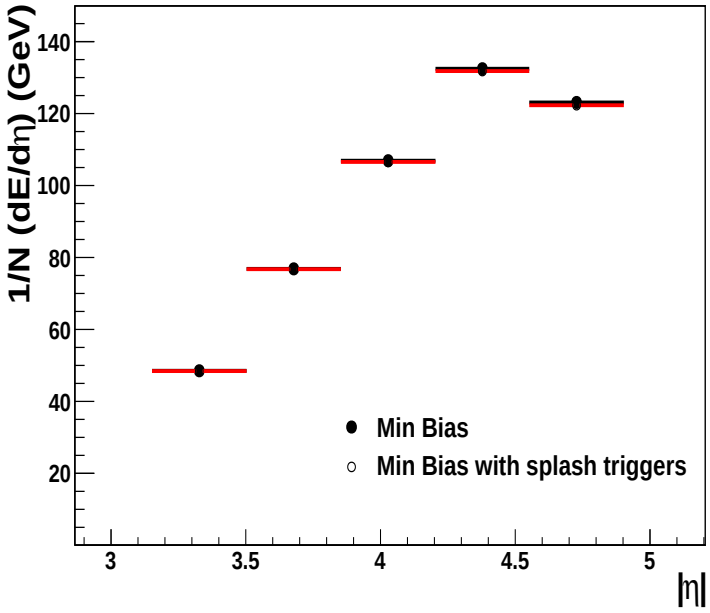

Figure 7.3: Impact of splash trigger on the energy flow measurement - Here, the influence of the splash triggers for minimum bias events, $\sqrt{s} = 7$ TeV, is shown. [171]

7.1.3 Vertex Selection

In previous QCD analyses (e.g. UE measurement [44]), the selection required two or more tracks associated to the primary vertex, thereby keeping the efficiency for event selection high and the fake-rate manageable. In this analysis, we impose a stricter requirement of five or more tracks, thereby reducing the vertex fake rate to zero [151] [170]. In Figure 7.4 we see that the difference in energy flow between the two different NDOF-limits for the vertex fit is less than 0.5% in the bin with the largest difference. We consider the systematic effect to be negligible.

7.1.4 HF Noise Cut-off

Figure 7.5 shows the electronic noise and other effects in the HF, using zero-bias events and minimum bias event samples [179]. To ensure the removal of all noise, a cross-check

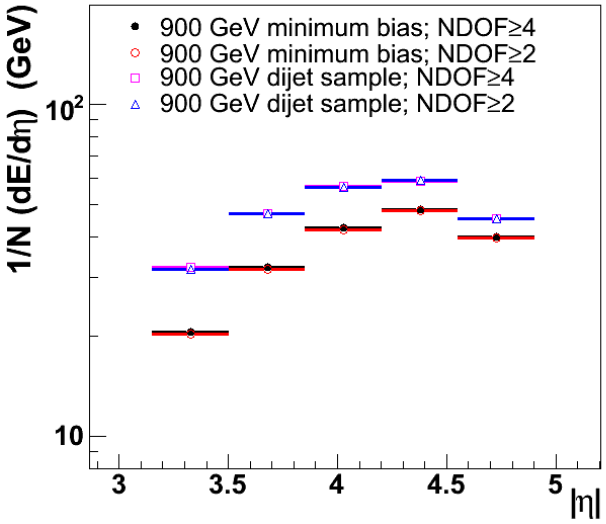

Figure 7.4: Impact of NDOF vertex criteria on energy flow - Variation of energy flow for different degrees-of-freedom requirements on the vertex fit, for minimum bias and dijet samples at $\sqrt{s} = 900$ GeV. [171]

is performed by analysing zero-bias samples and applying a 4 GeV limit; there is no energy deposit above the cut-off, signalling that the threshold is adequate.

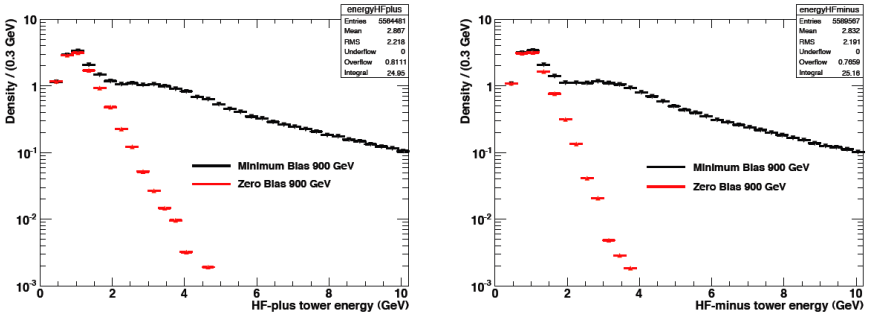

Figure 7.5: Noise and background HF towers - Shown for HF+ (left) and HF− (right), for the collision data sample (black) and the zero-bias sample (red) [179]. [171] [172]

7.1.5 Dead Material Effects

The HF calorimeter is shielded by the beam pipe, except at its two edges ($\eta \approx 2.9$ and $\eta \approx 5.2$). Additionally, at lower η ($2.9 < \eta < 3.1$), the calorimeter is shadowed

by the hadronic endcap calorimeter (HE) and some of the support structure of the electromagnetic endcap calorimeter (EE). These HF towers are in the shadow of the endcaps and serve to provide the full containment of hadronic showers of particles at $\eta \approx 3$.

We use a reduced HF-range (see Section 3.2.4.1) that is not shadowed by the endcap calorimeters, leaving the beam pipe as the main source of dead material effects. The other components, such as the sagging of the pipe and its support structure have a small effect, and are well modelled in the detector simulation. TOTEM T1 was not included in the simulation and the effect from its supporting structure is neglected as it is not in the path of the HF. Secondary scattering effects are neglected.

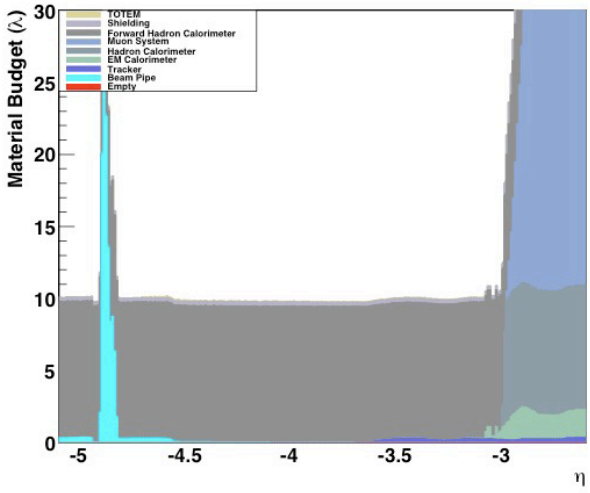

Figure 7.6: Distribution of material budget in front of the HF calorimeter - 10λ is the radiation length of the HF calorimeter, seen in grey. The cyan-coloured peak at $\eta \approx 4.9$ is dead material from the beam-pipe[184] [185].

Figure 7.6 shows the dead material in front of the HF calorimeter, with the beam pipe being the largest, or the most visible [184] [185]. The presence of the beam pipe around $\eta \approx 4.9$ explains the systematic migration effect seen in Chapter 6.3.1. It affects that particular region of the detector through the absorption and multiple-scattering of particles.

Figure 7.7 shows the effect of varying the material description in the detector simulation [184]. The blue and red lines show the energy flow distributions, when the

material budget is 0.5 – 2 times the existing description of dead material implemented in the detector simulation. The dashed-blue and dashed-red lines show that the energy flow varies by less than 1% when all the material except for the beam pipe is increased to 0.5 – 2 times the existing description. The material budget in other parts of the CMS detector, where one can measure it's effect from data using the tracker and electromagnetic calorimeter, is within 10%–20%. Considering an increase in density of the dead material by 50% (as done in Figure 7.7) is an over-estimation of the material budget error. Accordingly, we conclude that the main contribution comes from the beam pipe, and that all other sources of dead material have a negligible systematic effect on the energy flow.

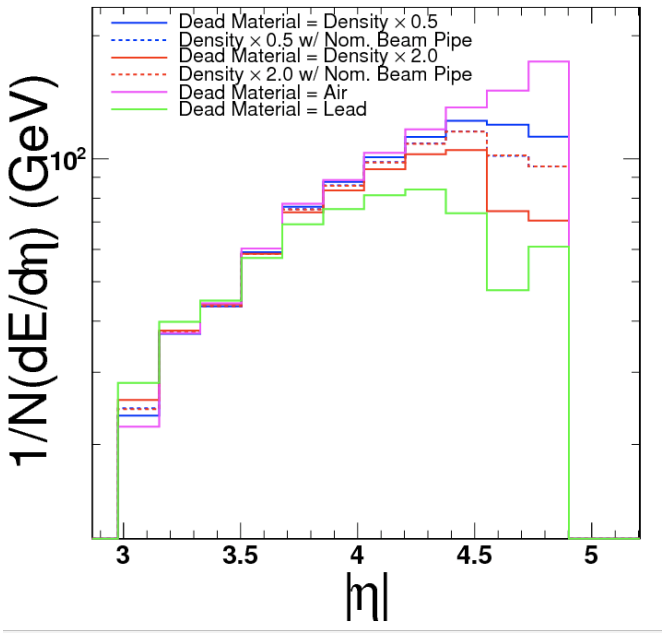

Figure 7.7: Variation of energy flow for different descriptions of dead material in the detector simulation - The green and pink lines indicate extreme situations, with lead and air being used as detector material. Blue and red represent an increase by a factor of 0.5–2, with and without the nominal beam pipe (dashed and solid) [184].

Figure 7.8 shows the distribution of energy flow as a function of *iphi* (see Section 3.2.4.1) in bins of η, for simulated events. We know that the xy–position of the vertex is displaced (tracker erroneously shifted in simulation [186]), resulting in the observed modulations. There are no additional modulations to indicate dead material effects.

The energy flow distributions as a function of *iphi* for data are shown in Figure 7.9,

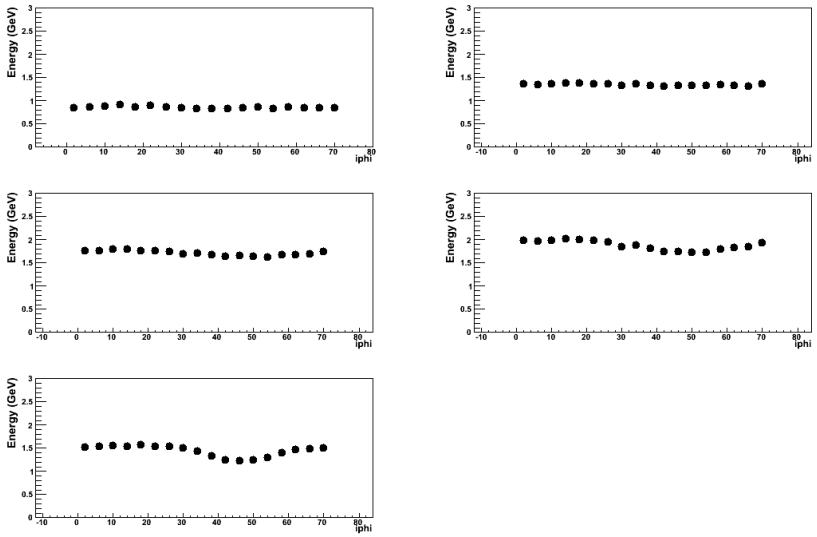

Figure 7.8: Energy flow as a function of *iphi*, using simulated events - For the five η ranges of the energy flow measurement, $\sqrt{s} = 900$ GeV. [171] [172]

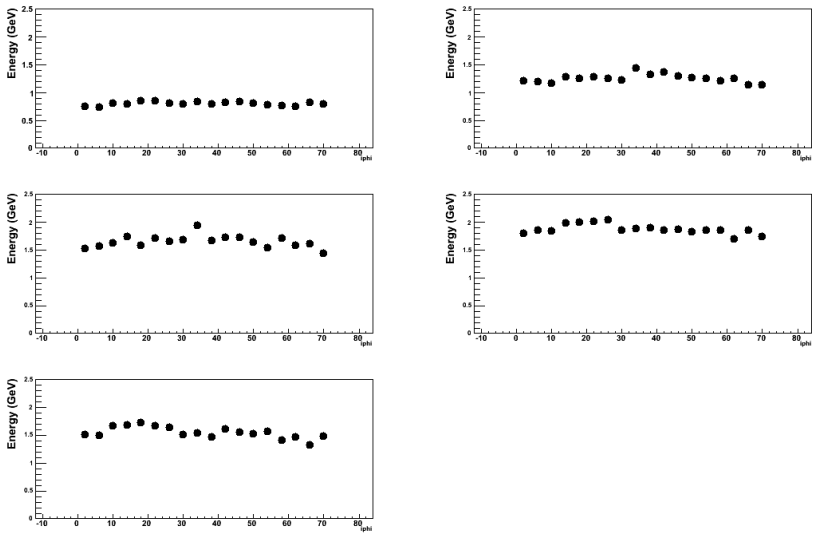

Figure 7.9: Energy flow as a function of *iphi*, using data - For the five η ranges of the energy flow measurement, $\sqrt{s} = 900$ GeV. [171] [172]

where we see random fluctuations corresponding to the lacking phi-calibration of the HF (covered by the energy scale uncertainty). However, there is no significant energy loss observed at a given *iphi*, suggesting that any dead material effects present are phi-symmetric.

7.1.6 Modelling of diffraction

MPI tunes to inclusive, non-diffractive data suffer from uncertainty on the diffractive component or background [187]. Both data and simulated samples retain a small percentage ($\approx$ 3% – 7%) of diffractive events, despite the non-single diffractive (NSD) selection criteria. PYTHIA 6.4, the generator used to calculate the correction factors, only includes a description of the soft component of diffraction; the different tunes estimate similar rates of diffractive events and do not vary by much in their treatment of the subject. Therefore, it is necessary to estimate the possible contributions and uncertainties due to incorrect estimates and modelling of hard diffraction, or lack thereof. PHOJET [86] models both the soft and hard diffractive component, and has been found to characterise measurements of single diffractive dissociation reasonably well [188]. A comparison of the energy flow predicted by the two event generators for two different scenarios, with and without diffraction helps, estimate the possible impact of proper diffractive modelling.

Figure 7.10 estimates a 4%–6% difference in measured energy flow due to diffraction, by comparing events without single diffraction (generator-level rejection of processes), and events with diffraction plus the NSD-trigger criteria (i.e., reduced diffraction). The contribution is small enough to be neglected. The ratio of the predictions from the generators, for the two scenarios (with and without) indicates that the models differ by very little, despite their dissimilarities. In other words, the absence of a proper treatment of diffraction by PYTHIA 6.4 does not influence the correction factor significantly; the effect of differing models is accounted for by the systematic uncertainty assigned to model dependence, and covered by the large systematic uncertainty (see following Sections).

7.1.7 Independence via Least Squares Method

Rather than the standard uses for χ^2/NDOF distributions, here it is adopted as a measure of the independence of the data from the correction procedure. The χ^2/NDOF

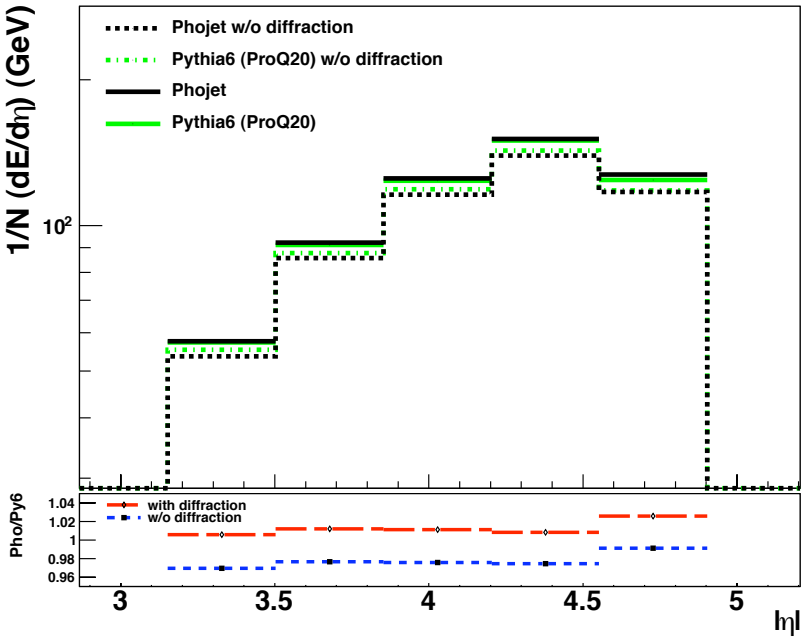

Figure 7.10: Contribution of the diffractive component to the energy flow - Energy flow with and without diffraction for minimum bias events at $\sqrt{s} = 7$ TeV, using two different models, PHOJET (black) and PYTHIA 6.4 PROQ20 (green). [171]

values for the data and the MC generators are listed in Table 7.1 (uncorrected) and Table 7.2 (corrected). χ^2 is calculated using the formula in Eq. 7.1, and NDOF refers to the number of degrees of freedom; in this case NDOF is the number of bins in the measurement, i.e. five degrees of freedom.

$$\chi^2 = \sum_{bins} (data - MC)^2 \tag{7.1}$$

It should be noted this procedure is not a true χ^2 calculation, as no uncertainties are included; statistical uncertainties are negligible, and the systematic uncertainties differ for uncorrected and corrected data. Thus, the values are not an indication of agreement between MC and data but can be used as a measure of bias introduced by the correction procedure. The order remains the same before and after corrections for most cases, signifying no bias was introduced.

Table 7.1: $\chi^2/ndof$ for comparisons between simulated MC predictions and uncorrected data. The fields marked with a dash indicate that the simulated MC sample does not exist. All uncertainties are neglected.

	Minimum Bias Data		Di-Jet Data	
MC prediction	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV
PYTHIA 6.4 D6T	9.1	195.4	59.7	1078.9
PYTHIA 6.4 P0	97.4	1350.0	200.1	5951.9
PYTHIA 6.4 PROQ20	29.4	338.6	4.0	160.1
PYTHIA 6.4 Z2	46.0	277.2	68.9	1686.5
PYTHIA 6.4 DW	3.7	-	251.9	-
PYTHIA 8	-	418.5	-	1352.2
PHOJET	-	-	-	6531.9

7.1.8 Other Effects

The contribution of beam-gas and non-interaction events is investigated by performing the same analysis for events with no beam-crossing. None of the 3.5 M events passes the selection criteria [171] [183].

Table 7.2: χ^2/NDOF for comparisons between hadron level MC predictions and corrected data. All uncertainties are neglected.

	Minimum Bias Data		Di-Jet Data	
MC prediction	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV
PYTHIA 6.4 D6T	55.0	1073.8	120.6	4124.6
PYTHIA 6.4 P0	414.3	5461.8	499.9	8508.0
PYTHIA 6.4 PROQ20	130.7	2886.6	23.1	321.5
PYTHIA 6.4 Z2	206.4	1483.1	179.3	2606.3
PYTHIA 6.4 DW	18.8	1747.9	596.5	1074.4
PYTHIA 8	277.8	13001.0	461.7	1419.0

7.2 Systematic Uncertainties

7.2.1 Energy Scale Uncertainty

The dominant systematic effect in the measurement of the forward energy flow is the global energy scale uncertainty of the HF calorimeters. The HF energy scale is determined using $Z \to e^+e^-$ events, data processed with a phi-symmetry calibration, and corrections to the electromagnetic calorimeter (ECAL)[189]. The typical energy of the electrons in $Z \to e^+e^-$ events spans over a large range, upto several hundred giga-electron volts (GeV), which is comparable to the typical energy of the particles measured by the HF; the e/π ratio is determined during beam-tests. Therefore, the energy scale can be applied for all particles detected by the calorimeter. The method outlined provides the new energy scale uncertainty of 10% [189]. This uncertainty is applied directly to the energy flow distributions and is the same for all bins in this study. With an improved measurement of the energy scale of the HF with increased amounts of data and tracking information, the overall uncertainty for this measurement could be reduced.

7.2.2 Primary Vertex z-position

The η is defined with respect to the $(0, 0, 0)$ point in the CMS reference system. For events with a primary vertex different from the $(0, 0, 0)$ point, the distributions of measured variables as a function of η are shifted. To evaluate the influence of this

effect, the energy flow is calculated for events with a primary vertex fulfilling $|z| < 4$ cm, $4 < |z| < 9$ cm and $9 < |z| < 15$ cm separately. The largest difference for the energy flow is found to be 1% [183] [184].

7.2.3 Channel-by-channel Miscalibration

To examine the effect of channel-by-channel miscalibration in the HF calorimeters, the response for each channel is randomly varied by $\approx 15\%$ [183] [184]. They are varied only once during each systematic study. The resulting energy flow is shifted by less than 1% in each of the cases.

7.2.4 Noise Threshold

Though the noise threshold is well determined using zero bias events, Figure 7.5 indicates that some energy remains beyond the 4 GeV limit, for one side of the HF. Since an increase in the threshold by 0.5 GeV would remove the remaining noise, the effect of such a variation is estimated. Similarly, a known feature of the HF-simulation is that a discrepancy between MC and data is observed (poor description of noise) for a threshold lower than 3.5 GeV. To gauge the possible influence of a change in the minimum energy-threshold criteria, the limit for the towers is varied by ± 0.5 GeV. The result (2% change) is included as a global systematic uncertainty [171] [172] [183].

7.2.5 Photomultiplier Hits

Events with particles hitting the read-out photomultipliers cause large signals in the HF calorimeter towers and are removed from the analysis. The rejection criteria are based on comparisons of the energy measurement in a given HF calorimeter cell with those in adjacent cells. The possible remaining noise from photomultiplier hits are estimated by using an alternative rejection algorithms. The resulting change in the energy flow is smaller than 3% [183] [184].

7.2.6 Non-linearity Effects

The non-linear response of the HF was investigated with test-beam data and its behaviour is properly simulated in the MC [176]. The only remaining systematic effect resulting from the non-linearity of the HF comes from the difference in the particle

spectra. Variations in particle spectra between the models produce disparate energy spectra, and thus dissimilar average energy flow due to the sub-detector's non-linear response. As long as the non-linearity of the simulation is well-modelled, this effect is covered by the model dependence of the correction factor (Chapter 7.2.9).

7.2.7 Non-uniformity Effects

Due to the non-uniformity arising from differences between the geometrical description of the detector in the simulation and the actual detector, we apply a set of corrections (Table 5.3), as described in Section 5.2.1.2. The uncertainty from the application of these corrections ($correction[i\eta]$) is determined using the following equation:

$$Uncertainty = 0.3 \cdot (1 - correction\,[i\eta]) \tag{7.2}$$

This leads to an uncertainty of $\approx 3\%$ in the highest η bin and is negligible for the lowest measurement bin. These values are a conservative estimate which reflect the difference between fibre counts in different HF wedges and some uncertainty in the placement of these wedges with respect to each other during the assembly.

7.2.8 Fibre Response

Discrepancies between data and simulation for low-energy particles have been observed [186] [190], indicating that the response from short fibres for low energetic particles is not correctly simulated. Moreover, the HF L-fibre sections are validated better with $Z \rightarrow e^+e^-$ events, than S-fibres due to the nature and depth of EM-cascades. The concern is that the dominant difference in particle composition between the different tunes may be related to low-energy π^0 - the overlap of many such low-p_T particles would result in significant non-linearity effects [190].

The HF energy scale uncertainty includes an energy scale calibration for both L and S sections separately. Additionally, low energy photons from π^0 decays produce signals primarily in the L-fibres of the HF; the impact of the modelling of the S-fibres have a low impact on the analysis of energy flow [176] [186]. Therefore, the possible discrepancies due to the S-fibre sections are absorbed by the uncertainty of the HF energy scale.

Nonetheless, it is prudent to assign an uncertainty until the issue is further investigated. The maximum deviation between the models in Figures 5.5–5.6 (Chapter 5.2.1.1) is adopted as an estimate of the uncertainty, for each of the measurements. Determined by repeating the analyses using long fibres exclusively, a 3% and 9% uncertainty was assigned for the minimum bias analyses. For the dijet analyses the uncertainty was found to be between 6% and 10%, except for the two highest rapidity bins at $\sqrt{s} = 900$ GeV where the uncertainty extends to 13% and 18%. Table 7.3 lists the uncertainties for each bin in the different measurements.

Table 7.3: Uncertainty attributed to the various measurements due to imperfect modelling of the short-fibre response in the simulation, for each bin.

	Minimum Bias		Di-jet	
Pseudorapidity range	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV
$3.152 < \eta < 3.503$	3.0	3.0	6.0	6.0
$3.503 < \eta < 3.853$	3.7	3.5	7.0	6.0
$3.853 < \eta < 4.204$	4.4	4.0	10.5	6.5
$4.204 < \eta < 4.552$	5.5	4.5	13.0	6.0
$4.552 < \eta < 4.903$	8.8	6.0	18.0	8.0
Range	3%-9%	3%-6%	6% - 18%	6%-8%

7.2.9 Model Dependence

The model dependence of the correction factor is evaluated using the PYTHIA 6.4 tunes that contributed to the correction procedure. Employing models which do not describe data prior to the correction procedure is unreasonable and the simulated predictions with PYTHIA 6.4 covers a large spread of distributions from other models. Moreover, the tunes were in reasonable agreement with the measured data prior to corrections applied. Therefore, it seems justified to only consider the PYTHIA 6.4 predictions.

The model dependent systematic uncertainties represent the maximum variation in the bin-by-bin correction factors calculated by using several different MC samples. The uncertainties are applied symmetrically around the data points, and vary between 4% (0.9 TeV min bias) and 17% (7 TeV dijets). The statistical uncertainty on the correction factors (and model dependence) is negligible. The fluctuations observed (illustrated in

Figures 7.11 and 7.12) are due to differences in particle spectra exaggerated by the hadron-level selection criteria (i.e. charged particle within $3.9 < \eta < 4.4$). The exact values are given in Table 7.4.

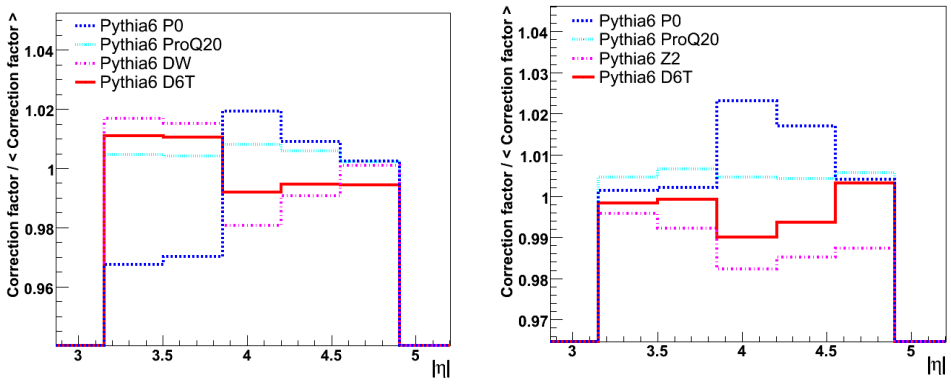

Figure 7.11: Model dependence of the correction factor, for minimum bias events - Left: $\sqrt{s} = 900$ GeV. Right: $\sqrt{s} = 7$ TeV [171]

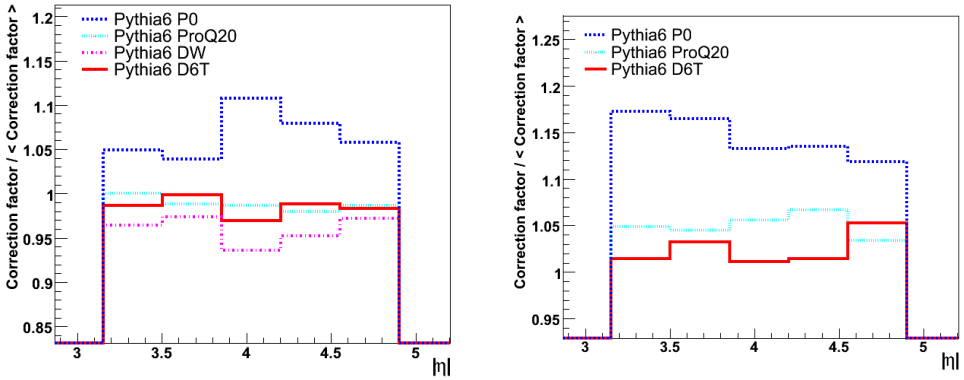

Figure 7.12: Model dependence of the correction factor, for dijet events - Left: $\sqrt{s} = 900$ GeV. Right: $\sqrt{s} = 7$ TeV. [171]

7.2.10 Jet Energy Scale Uncertainty

In the case of the dijet systems, an additional uncertainty arises from the jet energy scale (JES). Even though identical kinematic selection criteria were employed for hadron-level and detector-level jets, the samples still differ due to the jet energy resolution in the detector. The uncertainty on the JES at $\sqrt{s} = 7$ TeV is approximated at 10% [191].

Table 7.4: Model uncertainty for each bin, for the various measurements.

	Minimum Bias		Di-jet	
Pseudorapidity range	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV
$3.152 < \eta < 3.503$	3.2	0.5	5.0	17.0
$3.503 < \eta < 3.853$	3.0	0.8	4.0	16.4
$3.853 < \eta < 4.204$	2.0	2.3	10.5	13.0
$4.204 < \eta < 4.552$	1.0	1.7	8.0	13.5
$4.552 < \eta < 4.903$	0.6	1.4	5.7	12.0
Range	1%-3%	1%-2%	4% - 11%	12%-17%

Here, the jet's transverse momentum is shifted by the uncertainty on the JES, producing three different energy flow distributions in the HF (p_T = 18 GeV, 20 GeV and 22 GeV), within 2% of each other.

7.2.11 Total uncertainty

Often, a physics effect can impact several uncertainty estimations, thereby creating a correlation between them. For example, the particle spectra of an event ties in the energy scale and non-linear response of the calorimeter as well as the model dependence of the correction factors. In such instances, it is often impossible to disentangle or isolate the contribution of the effect to the individual uncertainties. To prevent an overestimation of the overall uncertainty due to possible double-counting, a conservative approximation by combining the individual estimates in quadrature is preferred.

Table 7.5 summarises the principle systematic effects. The uncertainties are added in quadrature and applied bin-by-bin. This leads to a global systematic uncertainty for each measurement, shown in Table 7.6. The dominant uncertainty is the energy scale for minimum bias events, and the short-fibre response and model dependence for dijet events. Possible ways to improve the precision of the measurement are discussed in Chapter 9.

Table 7.5: Largest systematic uncertainties affecting the measurement and its correction.

Effect	Minimum Bias	Dijets
HF energy scale	10%	10%
Non-uniformity	0%-3%	0%-3%
Short-fibre response, $\sqrt{s} = 0.9$ resp. 7 TeV	3%-9% / 3%-6%	6%-18% / 6%-8%
Model dependence, $\sqrt{s} = 0.9$ resp. 7 TeV	1%-3% / 1%-2%	4%-11% / 12%-17%
Jet energy scale, $\sqrt{s} = 0.9$ resp. 7 TeV	-	2% / 2%
Other effects	$\leq 1\%$	$\leq 1\%$

Table 7.6: Systematic uncertainty affecting the correction for each measurement, overall and for each bin.

	Minimum Bias		Dijet	
Pseudorapidity range	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 900$ GeV	$\sqrt{s} = 7$ TeV
$3.152 < \eta < 3.503$	11.6%	11.2%	13.5%	21.1%
$3.503 < \eta < 3.853$	11.8%	11.4%	13.6%	20.6%
$3.853 < \eta < 4.204$	11.9%	11.8%	18.5%	18.3%
$4.204 < \eta < 4.552$	12.3%	12.0%	18.9%	18.5%
$4.552 < \eta < 4.903$	14.2%	12.8%	22.0%	18.4%
Range	11%-14%	11%-13%	13% - 22%	18%-21%

8

Discussion of Results

This chapter contains the results of measurements of the energy flow in minimum bias and dijet events for the two centre-of-mass energies, $\sqrt{s} = 0.9$ TeV and 7 TeV. The data, corrected for detector effects, are compared to Monte Carlo predictions in an attempt to quantify which QCD effects impact the forward energy flow the most. The transverse energy flow is calculated to stack this measurement up against previous measurements at ep and $p\bar{p}$ colliders. Finally, the chapter concludes with a summary and an outlook for future studies of the transverse energy flow in pp collisions.

The energy flow is measured with the CMS hadronic forward calorimeters at large pseudorapidities ($3.15 < |\eta| < 4.9$), and corrected to the hadron level. The corrected data and the comparisons to Monte Carlo predictions are shown in Figures 8.1–8.5. The systematic uncertainties are indicated as error bars; they are correlated between η-bins. The statistical uncertainties are negligible and therefore not shown. The plots demonstrate three key traits observed in the measurements: a development in energy flow as a function of pseudorapidity; an increase of the same for higher centre-of-mass energies; a boost in energy for events with a hard scale set by a dijet system.

Figure 8.1 (left) contrasts various PYTHIA 6.4 tunes, holding them up to data. The Z2 and AMBT1 tunes, based on LHC data, and the older D6T tune are similar in their predictions, underestimating the activity in the forward region. Other parameter settings, such as PROQ20 and DW, tuned to TEVATRON data, or newer ones like P11 which use recently-developed MPI models, also fail to adequately describe the data; the variation of the tunes is of the order of $\sim 10 - 20\%$.

Figure 8.1 (right) accentuates the differences between assorted event generators and their model philosophies through a comparison to minimum bias data. PYTHIA 8 predictions obtained using the default version of the generator (i.e., without prior tuning to recent data) misestimates (far too small) the energy flow. The PYTHIA 6.4 tunes are presented as a band, which is constructed from the maximum and minimum values of the predictions in each bin. While the PYTHIA 8 predictions are always within the tune uncertainty band, PYTHIA 6.4 performs better despite the lack of hard diffraction modelling, signifying tuning effects largely outweigh diffractive contributions.

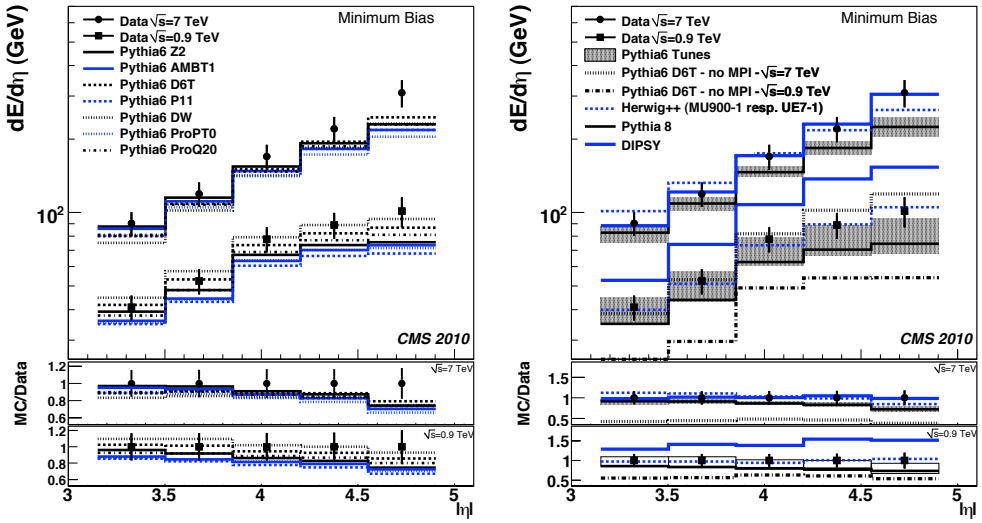

Figure 8.1: Energy flow measured in the minimum bias sample, data corrected to hadron level, $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, compared to pp-physics generators - The measured data corrected to hadron level are shown as points, the histograms are the predictions from PYTHIA 6.4 using various tunes (left) and PYTHIA 8, HERWIG++ and DIPSY (right). The coloured band is constructed from the maximum and minimum values of the PYTHIA 6.4 predictions in each bin. The error bars on the data represent the systematic uncertainties. Statistical uncertainties are negligible. [192] [181]

HERWIG++2.5, featuring angular-ordered parton showers, cluster fragmentation, and $\sqrt{s}$-specific tunes, describes the data well. Its parameters for the UE and colour reconnections were tuned to UE and minimum-bias data at $\sqrt{s} = 0.9$ TeV and to UE data at $\sqrt{s} = 7$ TeV. The agreement after tuning indicates strong correlations between central and forward UE activity. Without prior tuning to recent data, DIPSY reproduces the shape and magnitude of the minimum-bias data for $\sqrt{s} = 7$ TeV, but exhibits up to 50% surplus activity for $\sqrt{s} = 0.9$ TeV. The model's performance at the

higher centre-of-mass energy is promising; set to experimental data, the dipole picture of BFKL evolution emphasised by DIPSY could shed new light on evolution dynamics in minimum bias events.

Each of the aforementioned generators include MPI modelling within their framework. Evaluating the significance of MPI in the forward region using PYTHIA 6.4-D6T-without-MPI resulted in predictions 40% lower than the minimum bias measurements at both centre-of-mass energies. To ascertain that hadronisation effects and increased parton showering would not make up for the lack of MPI activity, the relevant parameters were modified; augmenting parton shower scales by a factor of four compensates for $\sim$ 5% of the difference, while comparing parton-level to corrected data estimates hadronisation effects account for no more than 20% of the energy [192] [193]. The studies illustrate the necessity of MPI constructs for physics in the forward region.

Figure 8.2 underscores the similarities between cosmic ray air showers and forward physics in pp collisions. Distributions derived from EPOS, QGSJET and SYBILL, event generators used in cosmic-ray air shower investigations, fair better than most obtained from standard collision-physics event generators [194]. The manner in which these models take contributions from both soft and hard parton dynamics into consideration exposes the deficiencies in current forward physics modelling.

The juxtaposition of energy flow data and simulations from PYTHIA 6.4 (Figure 8.3-left and Figure 8.4-left), assorted collision event generators (Figure 8.3-right and Figure 8.4-right), and cosmic-ray Monte Carlos (Figure 8.5) for dijet events reveals the same behaviour and properties as observed in minimum bias events. In data, the energy flow increases with $|\eta|$ and centre-of-mass energy. HERWIG++ matches data, DIPSY performs well at $\sqrt{s} = 7$ TeV, PYTHIA 6.4 encapsulates the data, PYTHIA 8 does not fair well, predictions without MPI underestimate activity significantly, and cosmic-ray event generators outperform standard event generators.

However, some distinguishing features emanate from the comparisons. Not only is the average energy in dijet events higher than in minimum bias, the ratio of energy flow between the two centre-of-mass energies is much greater; three to four times larger in dijet systems as opposed to two-three times in minimum bias topologies. In MC, one particular tune from PYTHIA 6.4 (Pro-Q20) duplicates attributes of the data well, PYTHIA 8 describes a part of the data at $\sqrt{s} = 7$ TeV, and two of the cosmic-ray MCs (QGSJET II and SIBYLL) show a larger deviation at the lowest $|\eta|$-bins for $\sqrt{s} = 0.9$ TeV.

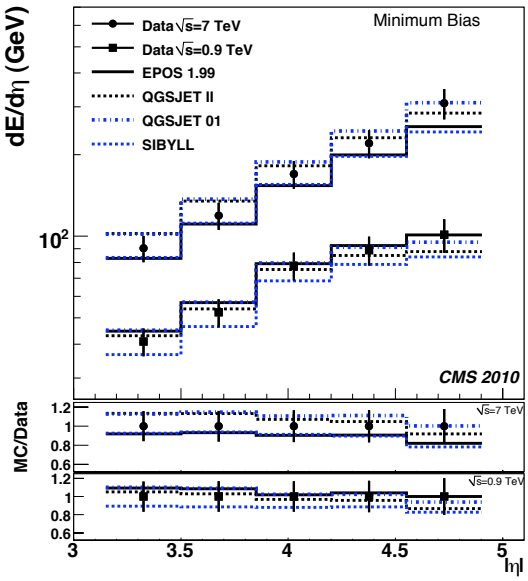

Figure 8.2: Energy flow measured in the minimum bias sample, data corrected to hadron level, $\sqrt{s} = 0.9\ TeV$ and $\sqrt{s} = 7\ TeV$, compared to cosmic-ray Monte Carlo generators - Energy flow in the minimum bias sample, as a function of η for $\sqrt{s} = 0.9\ TeV$ and $\sqrt{s} = 7\ TeV$. The measured data corrected to hadron level are shown as points, the histograms are the predictions from different event generators for cosmic ray physics. The error bars on the data represent the systematic uncertainties. Statistical uncertainties are considered negligible. [192] [181]

Additionally, CASCADE, an event generator based on the CCFM evolution equation and lacking a MPI-paradigm, offers an alternative approach to parton-showers. It covers a part of the difference between D6T- no-MPI and data. Nonetheless, the preponderance of secondary and tertiary collisions (semi-hard MPI) is evident, as indicated by the additional energy flow activity in dijet events. Lastly, the overall agreement with data is slightly better than in minimum bias topologies; the primary interactions in collision events are well-modelled.

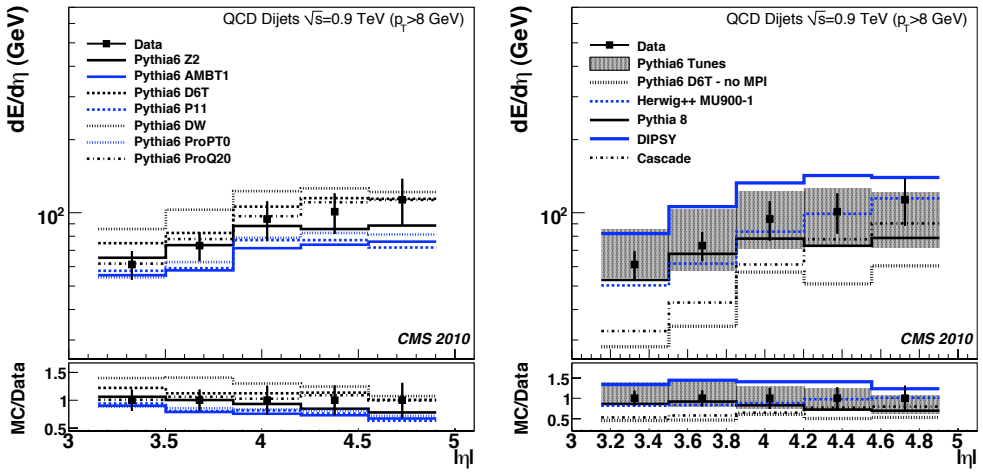

Figure 8.3: Energy flow measured in the dijet sample, data corrected to hadron level, $\sqrt{s}$ = 0.9 TeV, compared to hadron-level predictions made pp-physics Monte Carlo generators - Energy flow in the dijet sample as a function of η for $\sqrt{s}$ = 0.9 TeV. The measured data corrected to hadron level are shown as points, the histograms are the predictions from PYTHIA 6.4 using various tunes (left) and PYTHIA 6.4, PYTHIA 8, HERWIG++, CASCADE and DIPSY (right). The coloured band is constructed from the maximum and minimum values of the PYTHIA 6.4 predictions in each bin. The error bars on the data represent the systematic uncertainties. Statistical uncertainties are considered negligible. [192] [181]

Previous experiments at HERA [32] and SPS [195] have measured transverse energy flow. In order to compare this measurement to equivalent investigations at ep- and $p\bar{p}$-colliders, the transverse energy flow must be estimated from the average energy flow. Values obtained using Eq. 8.1, where θ was the mean of the bin, produced the curves shown in Figures 8.6–8.7.

$$\langle E_T \rangle \approx \langle E \rangle \sin\theta = \langle E \rangle \sin(2 \cdot \arctan(e^{-\eta})) \tag{8.1}$$

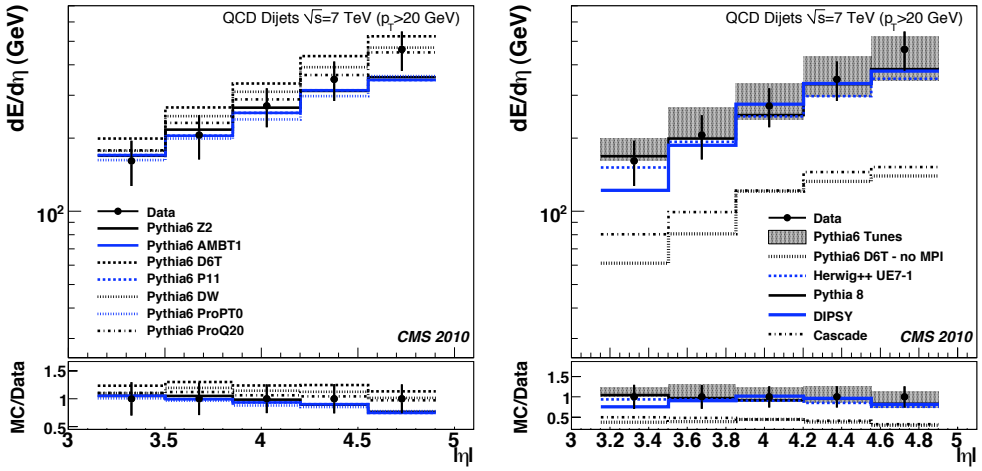

Figure 8.4: Energy flow measured in the dijet sample, data corrected to hadron level, $\sqrt{s} = 7$ TeV, compared to hadron-level predictions made pp-physics Monte Carlo generators - Energy flow in the dijet sample as a function of η for $\sqrt{s} = 7$ TeV. The measured data corrected to hadron level are shown as points, the histograms are the predictions from PYTHIA 6.4 using various tunes (left) and PYTHIA 6.4, PYTHIA 8, HERWIG++, CASCADE and DIPSY (right). The error bars on the data represent the systematic uncertainties. Statistical uncertainties are considered negligible. [192] [181]

Minimum bias events feature a constant transverse energy flow, $dE_T/d\eta \sim 3$ GeV (6 GeV) at $\sqrt{s} = 0.9$ TeV (7 TeV), compatible with the steep slope observed for $dE/d\eta$ distributions. $dE_T/d\eta$ values determined from Figure 8.7 for dijet events reveal that the gradient is consistent with p_T- or virtuality- (Q^2-) ordered parton-evolution calculations. That is, the lowest transverse momentum is in the direction of the proton, converging towards the E_T-limit of the hard scatter in the central region (i.e., DGLAP-like). Extrapolating the given numbers would yield $\sim$ 8 GeV ($\sim$ 20 GeV) at the collision-point ($\approx \eta = 0$) for $\sqrt{s} = 0.9$ TeV (7 TeV), accordant with the p_T-cut-off for the dijet systems. The theoretical predictions with MPI are in agreement with data for minimum bias events, and overestimate the energy flow substantially in dijet events, as seen in the energy flow distributions.

The average transverse energy in minimum bias events, $dE_T/d\eta \sim 2.4$ GeV at $\eta = 0$, previously measured in $p\overline{p}$-collisions [195] is comparable to the calculated value $dE_T/d\eta \sim 3$ GeV at $\eta = 3.5$ for $\sqrt{s} = 0.9$ TeV, given the flat progression seen in Figure 8.6. However, the calculation for $\sqrt{s} = 7$ TeV invalidates the assumed quadratic logarithmic dependence on $\sqrt{s}$ of the average transverse energy (Eq. 8.2) reported by

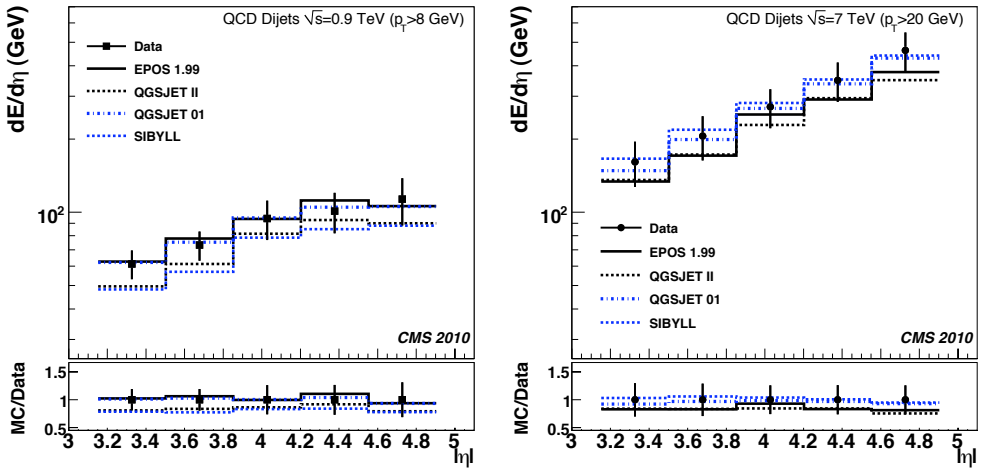

Figure 8.5: Energy flow measured in the dijet sample, data corrected to hadron level, $\sqrt{s} = 0.9$ TeV and $\sqrt{s} = 7$ TeV, compared to hadron-level predictions made cosmic-ray physics Monte Carlo generators - Energy flow in the dijet sample as a function of η for $\sqrt{s} = 0.9$ TeV (left) and $\sqrt{s} = 7$ TeV (right). The measured data corrected to hadron level are shown as points. The coloured band is constructed from the maximum and minimum values of the PYTHIA 6.4 predictions in each bin. The error bars on the data represent the systematic uncertainties. Statistical uncertainties are considered negligible. [192] [181]

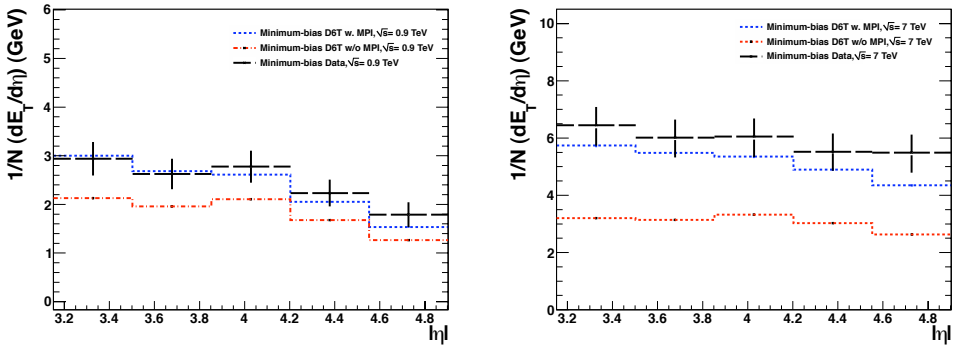

Figure 8.6: Transverse Energy Flow in Minimum Bias Events - Transverse energy flow estimated from average energy flow data using Eq. 8.1, $\sqrt{s} = 0.9$ TeV (left) and $\sqrt{s} = 7$ TeV (right). The error bars represent the systematic uncertainty for energy propagated to the transverse energy flow. Statistical uncertainties are neglected.

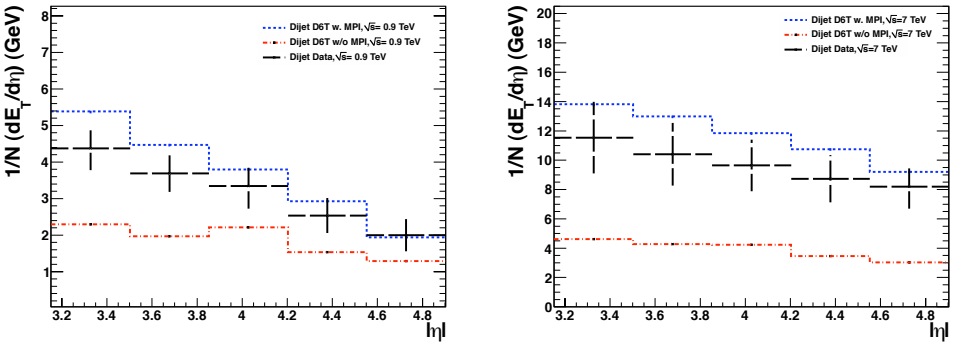

Figure 8.7: Transverse Energy Flow in Dijet Events - Transverse energy flow estimated from average energy flow data using Eq. 8.1, $\sqrt{s} = 0.9$ TeV (left) and $\sqrt{s} = 7$ TeV (right). The error bars represent the systematic uncertainty for energy propagated to the transverse energy flow. Statistical uncertainties are neglected.

the UA1 collaboration.

$$\frac{d(\sum E_T)}{d\eta}(GeV) = 0.80 - 0.075\ln(\sqrt{s}/GeV) + 0.051(\ln(\sqrt{s}/GeV))^2 \qquad (8.2)$$

The comparison of the average transverse energy flow calculated for dijet events at CMS and that measured in deep-inelastic scattering events (DIS) [32] at similar x and Q^2 is shown in Table 8.1. No corresponding measurement exists for the 7 TeV calculation. In DIS-events, where the contributions from MPIs are negligible, the transverse energy flow is much smaller than for pp-collisions.

Table 8.1: Average transverse momentum and their corresponding kinematic values (x, $Q^2 = 4p_T^2$, η) for two separate measurements, one at CMS and the other in DIS events at H1 [32]. Each of the values is an approximation.

Measurement	Q^2 (GeV^2)	x	η	$dE_T/d\eta$
CMS di-jet	250	0.02	3.3	4
H1 DIS	250	0.01	~ 3	2

In summary, while the shape of the energy flow distributions is reasonably well-reproduced by the event generators, few described the magnitude of measured activity. Adjusting the UE-related parameters to experimental data trumped all other considerations. Furthermore, it has been shown that MPIs produce a significant amount of

the hadronic energy flow. Therefore, the measured energy flow can be used for further constraints of the models with multiple parton interactions. The exact data points along with their corresponding uncertainty are listed in Tables 8.2 and 8.3. Cosmic-ray event simulations were shown to entail a more complete modelling of soft and hard parton dynamics for boosted collisions. Finally, calculations of transverse energy flow are in agreement with previous comparable measurements and expectations. These measurements form a good basis for a more extensive study of $dE_T/d\eta$ with particle flow objects, spanning the entire pseudorapidity range of CMS.

Table 8.2: Corrected energy flow ($1/N(dE/d\eta)$) and systematic uncertainties (δ_{sys}) for the minimum bias measurements. The units of the data and the uncertainties are in GeV. The statistical errors are in all bins less than 0.1%, and therefore not listed.

	Minimum Bias Data			
	$\sqrt{s} = 900$ GeV		$\sqrt{s} = 7$ TeV	
η bin range	$1/N(dE/d\eta)$	$\pm\delta_{sys}$	$1/N(dE/d\eta)$	$\pm\delta_{sys}$
3.152 - 3.503	40.9	4.8	90.2	10.1
3.503 - 3.853	52.4	6.2	119.3	13.6
3.853 - 4.204	77.8	9.2	169.7	20.0
4.204 - 4.552	88.8	10.9	220.3	26.4
4.552 - 4.903	101.2	14.4	309.6	39.5

Table 8.3: Corrected energy flow ($1/N(dE/d\eta)$) and systematic uncertainties (δ_{sys}) for the di-jet measurements. The units of the data and the uncertainties are in GeV. The statistical errors are in all bins less than 0.1%, and therefore not listed.

	Di-Jet Data			
	$\sqrt{s} = 900$ GeV		$\sqrt{s} = 7$ TeV	
η bin range	$1/N(dE/d\eta)$	$\pm\delta_{sys}$	$1/N(dE/d\eta)$	$\pm\delta_{sys}$
3.152 - 3.503	61.3	8.3	160.9	34.0
3.503 - 3.853	73.2	10.0	205.7	42.4
3.853 - 4.204	94.2	17.4	270.9	49.4
4.204 - 4.552	101.0	19.1	348.3	64.3
4.552 - 4.903	113.1	24.9	463.1	85.0

9

Coda: In Through the Out Door

Energy flow, $dE/d\eta$, has been studied in proton-proton collisions at the LHC, for two centre-of-mass energies, $\sqrt{s}$ = 0.9 TeV and 7 TeV. The measurements are made in the pseudorapidity range $3.15 < |\eta| < 4.9$, for both minimum-bias events and events with at least two high-momentum jets, using the CMS detector. Prior to the analysis presented in this thesis, an investigation of the energy flow in this η-region has never been reported at hadron colliders.

9.1 Summary

The LHC is a hadron collider that provides a unique opportunity to study proton-proton collisions in a previously unexplored phase space. Quantum chromodynamics sets the formalism needed to characterise hadron-hadron collisions provided at such an accelerator. However, the scale of the interaction must be sufficiently large for the perturbative approach to be applicable. Given the impact of collinear effects and the fact that the hadron structure is non-perturbative in nature, pQCD is replaced by phenomenological models at the Λ_{QCD}-limit. These models are tuned to describe data from previous experiments, but are limited in their description of current measurements at the LHC. Calculations indicate an increase in the total cross section due to additional multi-parton and underlying event activity with respect to colliders such as the TEVATRON. Therefore, the extrapolation of existing tunes to larger centre-of-mass energies is uncertain, and often leads to sizeable discrepancies amongst the various predictions.

This measurement seeks to provide additional constraints to existing parameter sets of the underlying event by studying low-x physics. With an unprecedented coverage in pseudorapidity provided by forward calorimeters, the CMS detector allows for an investigation of UE models over a larger phase space than was possible in previous experiments. The Hadronic Forward calorimeter is a sub-detector which extends from $2.9 < |\eta| < 5.2$, thereby equipping CMS for investigations at large pseudorapidities; it is segmented into 13 rings in pseudorapidity and is completely ϕ-symmetric. However, a reduced range ($3.15 < |\eta| < 4.9$) is used to limit possible shower leakage from overlapping detectors and discrepancies between simulation-data comparisons. The full event-simulation chain includes accurate modelling of detector effects in the measurement region. Where relevant, additional factors adjust the simulated variables or data to better account for actual effects such as non-uniformity of the detector.

The observable chosen to investigate the forward region is energy flow, as it spans a larger phase-space region than available track-related variables, and describes the shape and magnitude of the event globally (i.e. independent of particle and shower type). The measurement is presented as a function of pseudorapidity; ten η-rings are combined into five bins by grouping two consecutive rings together. Two event topologies are investigated: minimum bias events and events with a central dijet system. Minimum bias is defined as non-single diffraction events with at least one good reconstructed vertex in a pp collision. Dijet events consist of at least two high-momentum jets within the central region of CMS ($|\eta_{jet}| < 2.5$), with the momentum scale of the jets set to comparable x-values for the two centre-of-mass energies. The jets are defined using the anti-k_T algorithm, with $p_{T,jet} > 8$ GeV ($p_{T,jet} > 20$ GeV) for $\sqrt{s} = 900$ GeV ($\sqrt{s} = 7$ TeV); relative and absolute corrections are applied to the jet momenta. There is less than 1% of pile up events in the data samples, and therefore no special treatment is required for the measurements.

A bin-by-bin correction method is used to unfold the data to hadron level. The correction factors for each of the measurements are calculated by taking the ratio of hadron-level predictions to detector-level simulations. Hadron-level distributions are generated with selection criteria similar to actual requirements in data. That is, emulating the detector-level trigger for NSD event selection by requiring at least one charged particle on either side of the interaction point ($3.9 < |\eta| < 4.4$). In the case of the dijet analysis, the kinematic selection of the dijet system (p_T, η,$\Delta\phi$) was identical

for both data and simulations. The four PYTHIA6.4 detector-level predictions (tunes) that best describe and encapsulate the uncorrected data are chosen for the correction factor, providing an averaged ratio.

Studies with single-particle and minimum bias samples show migration effects to have a negligible contribution to the measured energy flow. Furthermore, a systematic check via the least-squares method demonstrates that no bias is introduced by the correction method, ensuring the applicability of the calculated factors. However, the analysis suffers from large systematic uncertainties caused by the variations in the models used. Supplementary systematic studies estimate further uncertainties attributed to the modelling of the short-fibre response in simulations, dependence on the non-uniformity corrections applied, and the poor energy resolution of the calorimeter. The global systematic uncertainty assigned to the measurement is a result of the individual estimates added in quadrature.

The corrected data display three general trends - an increase as a function of pseudorapidity, an increase with centre-of-mass energy, and a dependence on the presence of the hard scale. The plots show that the shape of theoretical predictions from event generators reproduce these characteristics; the variation of energy flow with η and that the presence of the high momentum dijet system in the central region increases the deposited energy in the entire phase space. However, many of the distributions from the MC generators fail to predict the amount of energy flow in an event. Models which include a description of multi-parton interactions produce predictions closer to data than those without. Varying parameters which impact hadronisation and the amount of parton radiation cannot compensate for the lack of a working MPI model. Nonetheless, the large spread in the theoretical predictions indicates sensitivity of the measurement shows sensitivity to the variations amongst the tunes.

The PYTHIA generator has been used to highlight the differences between various MPI models, changes in impact parameters, and ISR and FSR contributions. While none of the PYTHIA6.4 tunes describe all four of the measurements, D6T and PROQ20 fair well in minimum-bias events and in dijet events, respectively. HERWIG++ features angular-ordered parton showers, the cluster hadronisation model, and colour-string reconnections; the new centre-of-mass dependent tunes are in agreement with the data. DIPSY and CASCADE explore alternative parton dynamics, BFKL and CCFM respectively; predictions from DIPSY fail in the low-$\sqrt{s}$ scenarios and the absence of an MPI model

limits the validity of CASCADE. The best description of the measured energy flow comes from models based on cosmic-ray interactions.

Finally, the transverse energy flow is calculated, allowing a comparison with measurements in deep-inelastic scattering events (ep collisions) at HERA. As expected, the activity in pp collisions is significantly higher compared to ep collisions where MPI contributions are negligible.

9.2 Outlook

This analysis not only sets a baseline (reference) for future measurements at large pseudorapidities, but also provides the first experimental evidence of the forward region's sensitivity to underlying event activity and various QCD effects. The variations in shape and magnitude amongst the models is large and an important application of this measurement will be to further constrain UE parameters. With the first wave of CMS tunes emerging, a new parameter set using this measurement in conjunction with other analyses of QCD and UE structure (e.g. $\frac{dN}{d\eta}$) is expected. Moreover, this study affords us with the opportunity to characterise the scaling of UE activity as a function of centre-of-mass energy; additional measurements at $\sqrt{s} = 2.36$ TeV, and possibly higher energies (e.g., $\sqrt{s} = 2.7$ TeV, 8 TeV and 10 TeV) will produce an interesting constraint that could reduce uncertainties in future extrapolations. A natural extension of this analysis would be investigations of the energy flow's dependence on the central jet system, and a comparison between average correction factors and corrections obtained via the energy spectrum.

Of course, before embarking on further investigations of the UE, a necessary step would be to improve the precision of this measurement. Additional data will not affect the statistical uncertainty (negligible), but the large uncertainties due to systematic effects can be constrained. One of the dominant systematic uncertainties is the model dependence of the correction factor; further tuning the event generators or selecting tunes more compatible with the data, as well as re-weighting the distributions prior to applying unfolding methods could lead to a smaller model dependence. Also, an improvement in the geometry description removes the need for artificial adjustments to account for geometrical non-uniformity effects, thus reducing the systematic uncertainty associated with it. Another option is to extend the measurement over the entire

region covered by CMS using particle flow information. Exploiting the supplementary information available (e.g., better detector resolution, both positional and in energy), would lead to a reduced systematic uncertainty associated to calorimeter effects (e.g., energy scale, non-uniformity) and to probe very low-p_T activity. Lastly, the interpretation of the energy flow variable was complicated due to its natural dependence on pseudorapidity. Choosing a different set of variables such as transverse energy and possibly particle multiplicity would reveal more about the contributions to the underlying event and their inter-dependence.

Despite the numerous ways in which this study could be ameliorated, the four measurements have provided some insight into the activity at large pseudorapidities, as well as the issues that affect such an analysis at CMS. They have lead to a better understanding of activity and basic processes in hadron collisions at the LHC, demonstrating that the forward region is sensitive to both perturbative and non-perturbative QCD. In closing, I hope that the investigation presented in this thesis will set the foundations for further measurements in this interesting, unexplored region of phase space. This concludes my three-year bummel through the world of particle physics.

"A 'Bummel'," I explained, "I should describe as a journey, long or short, without an end; the only thing regulating it being the necessity of getting back within a given time to the point from which one started. Sometimes it is through busy streets, and sometimes through the fields and lanes; sometimes we can be spared for a few hours, and sometimes for a few days. But long or short, but here or there, our thoughts are ever on the running of the sand. We nod and smile to many as we pass; with some we stop and talk awhile; and with a few we walk a little way. We have been much interested, and often a little tired. But on the whole we have had a pleasant time, and are sorry when 'tis over." [196]

Taken from Three Men on the Bummel, written by Jerome K. Jerome

References

[1] Carlo Rovelli. **Naissance de la Science**. Lectures, 2004. 1

[2] E. Nagy. **Physique Subnuclèaire**. Lectures, 2007. 1

[3] D. Griffiths. *Introduction to Elementary Particles*. John Wiley and Sons, New York, 1987. 1, 6

[4] W. N. Cottingham and D. A. Greenwood. *An Introduction to the Standard Model of Particle Physics*. Cambridge University Press, 1998. 2

[5] T. Morii et al. *The Physics of the Standard Model and Beyond*. World Scientific Publishing, 2004. 2

[6] Jean Hladik and Michel Chrysos. *Introduction á la Relativitè Restreinte*. Dunod, 2006. 2

[7] I. J. R. Aitchison and A. J. H. Hey. *Gauge Theories in Particle Physics: From Relativistic Quantum Mechanics to QED*. The Institute of Physics, 2000. 2

[8] C. Cohen-Tannoudji et al. *Quantum Mechanics*. Hermann, 1997. 2

[9] B. F. Schutz. *A First Course in General Relativity*. Cambridge University Press, 1985. 2

[10] C. Amsler et al. **Prticle Data Group**. *Phys. Lett.*, **B667**, 2008. 2

[11] P. W. Higgs. **Broken symmetries, massless particles and gauge fields**. *Phys. Lett. 12*, page 132, 1964. 2

[12] P. W. Higgs. **Spontaneous Symmetry Breakdown without Massless Bosons**. *Phys. Lett.*, **145**:1156, 1966. 2

[13] F. Englert and R. Brout. **Broken symmetry and the mass of the gauge vector mesons**. *Phys. Rev. Lett.*, **13**:321, 1964. 2

[14] G.S. Guralnik, C. R. Hagen, and T. W. B. Kibble. **Global Conservation Laws and Massless Particles**. *Phys. Rev. Lett.*, **13**:585, 1964. 2

[15] E. Witten. **Dynamical breaking of supersymmetry**. *Nucl. Phys.*, **B188**:513, 1981. 2

[16] Romesh K. Kaul. **Gauge Hierarchy in a Supersymmetric Model**. *Phys. Lett.*, **B109**:19, 1982. 2

[17] Romesh K. Kaul. **SUPERSYMMETRIC SOLUTION OF GAUGE HIERARCHY PROBLEM**. *Pramana*, **19**:183, 1982. 2

[18] Gianfranco Bertone, Dan Hooper, and Joseph Silk. **Particle dark matter: Evidence, candidates and constraints**. *Phys.Rept.*, **405**:279–390, 2005. 2

[19] Steven Weinberg. **Implications of Dynamical Symmetry Breaking: An Addendum**. *Phys. Rev.*, **D19**:1277–1280, 1979. 2

[20] Leonard Susskind. **Lattice Models of Quark Confinement at High Temperature**. *Phys. Rev.*, **D20**:2610–2618, 1979. 2

[21] K. Lane. **Two Lectures on Technicolor**. *hep-ph0202255*, 2002. 2

[22] Christopher T. Hill and Elizabeth H. Simmons. **Strong dynamics and electroweak symmetry breaking**. *Phys. Rept.*, **381**:235–402, 2003. [Erratum-ibid.390:553-554,2004]. 2

[23] International Workshop on Strongly Coupled Gauge Theories. *Some Recent Results on Models of Dynamical Electroweak Symmetry Breaking*, 2006. 2

[24] N. Arkani-Hamed, A. G. Cohen, and H. Georgi. **Electroweak symmetry breaking from dimensional deconstruction**. *Phys. Lett.*, **B513**:232, 2001. 2

[25] J. Wess and B. Zumino. **Supergauge Transformations in Four-Dimensions**. *Nucl. Phys.*, **B70**:39, 1974. 2

[26] J. Wess and B. Zumino. **A Lagrangian model invariant under super- gauge transformations**. *Phys. Lett.*, **B49**:52, 1974. 2

[27] I Antoniadis. **A Possible new dimension at a few TeV**. *Phys. Lett.*, **B246**:377, 1990. 2

[28] N. Arkani-Hamed, S. Dimopoulous, and G. R. Dvali. **The Hierarchy problem and new dimensions at a millimeter**. *Phys. Lett.*, **B429**:263, 1998. 2

[29] Lisa Randall and Raman Sundrum. **A Large mass hierarchy from a small extra dimension**. *Phys.Rev.Lett.*, **83**:3370–3373, 1999. 2

[30] G. F. Giudice, R. Rattazzi, and J. D. Wells. **Quantum gravity and extra dimensions at high-energy colliders**. *Nucl. Phys.*, **B544**:3, 1999. 2

[31] C. Csaki, M. L. Graesser, and G. D. Kribs. **Radion Dynamics and Electroweak Physics**. *Phys. Rev.*, **D63**, 2001. 2

[32] The H1 Collaboration. **Measurements of transverse energy flow in deep inelastic-scattering at HERA**. *Eur. Phys. J.*, **C12**:595, 2000. xii, 3, 32, 33, 123, 126

[33] S. Weinberg. **A Model of Leptons**. *Phys. Rev. Lett. 19, 1264*, 1967. 5

[34] Abdus Salam. **Weak and Electromagnetic Interactions**. *Originally printed in Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden*, Stockholm 1968, 367-377*, 1968. 5

REFERENCES

[35] S. L. Glashow, J. Iliopoulos, and L. Maiani. **Weak Interactions with Lepton-Hadron Symmetry**. *Phys. Rev.*, **D2**:1285–1292, 1970. 5

[36] R. Keith Ellis, W. James Stirling, and B. R. Webber. *QCD and Collider Physics*, **8**. Cambridge University Press, 1996. 5, 6, 7, 8

[37] K. Nakamura et al. **PDG Particle Listings**. *J.Phys., G37:075021*, 2010. 5, 6

[38] M. Kaku. *Quantum Field Theory: A modern introduction*. Oxford University Press, 1993. 5

[39] J.D. Björken and S.D. Drell. *Relativistic Quantum Fields*. McGraw Hill, 1964. 5

[40] F. Yndurain. *Quantum Chromodynamics*. Springer-Verlag, 1983. 6, 7

[41] Yuri L. Dokshitzer et al. *Basics of Perturbative QCD*. Editions Frontieres, 1991. 6

[42] T. Muta. *Foundations of Quantum Chromodynamics*. World Scientific Publishing, 1987. 6, 7

[43] P. van Baal. *Confinement, Duality, and Non-Perturbative Aspects of QCD*. Kluwer Academic Publishers, 1997. 8

[44] The CMS Collaboration. **First Measurement of the Underlying Event Activity at the LHC with $\sqrt{s} =$ 0.9 TeV**. *Eur. Phys. J.*, **C70**:555, 2010. 9, 25, 28, 103

[45] The CDF Collaboration. **Studying the Underlying Event in Drell-Yan and High Transverse Momentum Jet Production at the Tevatron**. *Phys. Rev.*, **D**(82):034001, 2010. 9, 15

[46] J.W. Huston, W.J. Stirling, and J.M. Campbell. **Hard Interactions of Quarks and Gluons: A Primer for LHC Physics**. *Rept. Prog.Phys*, 2007. 10

[47] S. D. Drell and T.-M. Yan. **Partons and their applications at high energies**. *Ann. Phys.*, **66**:578, 1971. 10

[48] Daniel Stump, Joey Huston, Jon Pumplin, Wu-Ki Tung, H.L. Lai, et al. **Inclusive jet production, parton distributions, and the search for new physics**. *JHEP*, **0310**:046, 2003. 11

[49] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. **Parton Distributions for the LHC**. *Eur. Phys. J.*, **C**(63):189, 2009. 12

[50] F.D. Aaron et al. **Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA**. *JHEP*, **1001**:109, 2010. 12

[51] The H1 and ZEUS Collaboration. **QCD Fit HERAPDF1.5** [online]. 2010. H1prelim-10-142, ZEUS-prel-10-018 [cited 27.09.2011]. 12

[52] L. N. Lipatov. **The parton model and perturbation theory**. *Sov. J. Nucl. Phys.*, **20**:94–102, 1975. 12

[53] V. N. Gribov and L. N. Lipatov. **Deep inelastic** *ep scattering in perturbation theory*. *Sov. J. Nucl. Phys.*, **15**:438–450, 1972. 12

[54] Guido Altarelli and G. Parisi. **Asymptotic Freedom in Parton Language**. *Nucl. Phys.*, **B126**:298, 1977. 12, 13

[55] Yuri L. Dokshitzer. **Calculation of the Structure Functions for Deep Inelastic Scattering and $e+e-$ Annihilation by Perturbation Theory in Quantum Chromodynamics**. *Sov. Phys. JETP*, **46**:641–653, 1977. 12

[56] I. I. Balitsky and L. N. Lipatov. **The Pomeranchuk Singularity in Quantum Chromodynamics**. *Sov. J. Nucl. Phys.*, **28**:822=829, 1978. 13

[57] V. S. Fadin, E. A. Kuraev, and L. N. Lipatov. **On Pomeranchuk Singularity in Asymptotically Free Theories**. *Phys. Lett.*, **B60**:50, 1975. 13

[58] E. A. Kuraev, L. N. Lipatov, and Victor S. Fadin. **Multi - Reggeon Processes in the Yang-Mills Theory**. *Sov. Phys. JETP*, **44**:443–450, 1976. 13

[59] E. A. Kuraev, L. N. Lipatov, and Victor S. Fadin. **The Pomeranchuk Singularity in Nonabelian Gauge Theories**. *Sov. Phys. JETP*, **45**:199–204, 1977. 13

[60] T. Namsoo. *Three- and Four-Jet States in Photoproduction at HERA*. PhD thesis, University of Bristol, 2005. 13, 15, 31

[61] Marcello Ciafaloni. **Coherence Effects in Initial Jets at Small q^2/s**. *Nucl. Phys.*, **B296**:49, 1988. 14

[62] S. Catani, F. Fiorani, and G. Marchesini. **QCD Coherence in Initial State Radiation**. *Phys. Lett.*, **B234**:339, 1990. 14

[63] H. Jung and Gavin P. Salam. **Hadronic final state predictions from CCFM: the hadron-level Monte Carlo generator CASCADE**. *Eur. Phys. J.*, **C19**:351, 2001. 14, 16, 26

[64] Torbjorn Sjöstrand and Maria van Zijl. **A Multiple Interaction Model for the Event Structure in Hadron Collisions**. *Phys. Rev.*, **D36**:2019, 1987. 14, 19, 21

[65] Darin E. Acosta et al. **The underlying event in hard interactions at the Tevatron $\bar{p}p$ collider**. *Phys. Rev.*, **D70**:072002, 2004. 15, 28

[66] P. Bartalini and L. Fano, editors. *Studying the Underlying Event at CDF and the LHC*, **DESY-PROC-2009-06**. Proceedings of the First International Workshop on Multiple Partonic Interactions at the LHC-MPI08, 2008. 15, 25, 78

[67] The CDF Collaboration. **Charged Jet Evolution and the Underlying Event in Proton-Antiproton Collisions**, 2001. 15

[68] Ll. Marti. *Multiple Interactions in Photoproduction at HERA/H1*. PhD thesis, Universität Hamburg, 2008. 15

[69] Sakar Osman. *Multiple parton interactions in deep inelastic ep-scattering at HERA*. PhD thesis, Lund University, 2008. 15

[70] Hannes Jung, Lluis Marti-Magro, Tim Namsoo, and Sakar Osman. **Multiple interactions at HERA**. DESY-PROC-2009-02, 2009. 15

[71] J. Bartels, K. Borras, M Diehl, H. Jung, H. Abramowicz, et al., editors. *Elastic and diffractive scattering: Forward physics and QCD. Proceedings, International Conference, 12th Blois Workshop, EDS 2007, Hamburg, Germany, May 21-25, 2007*, number DESY-PROC-2007-02, 2007. 15, 31

[72] H. Jung. **QCD and Collider Physics IV**. Lectures, 2007. 15

[73] Livio Fano. **Multiple parton interactions, underlying event and forward physics at LHC**, 2007. Prepared for 12th International Conference on Elastic and Diffractive Scattering: Forward Physics and QCD, Hamburg, DESY, Germany, 21-25 May 2007. 16

[74] T. Sjostrand. **Monte Carlo Generators for the LHC**. Academic Training Lectures, 2005. 17

[75] G. Corcella et al. **HERWIG**6.5 **Release Note**. *hep-ph/0210213*, 2005. 16

[76] Torbjorn Sjöstrand, Stephen Mrenna, and Peter Skands. **PYTHIA 6.4: Physics and Manual**. *JHEP*, **05**:026, 2006. 16, 20, 22, 25, 78

[77] A. Donnachie and P. V. Landshoff. **Total cross-sections**. *Phys. Lett.*, **B296**:227–232, 1992. 16

[78] T. Sjostrand and Peter Z. Skands. **Multiple interactions and the structure of beam remnants**. *JHEP*, **0403**:053, 2004. 16, 20, 22

[79] G. Marchesini and B. R. Webber. **Monte Carlo Simulation of General Hard Processes with Coherent QCD Radiation**. *Nucl. Phys.*, **B**(310):461, 1988. 18

[80] B. R. Webber. **A QCD Model for Jet Fragmentation Including Soft Gluon Interference**. *Nucl. Phys.*, **B**(238):492, 1984. 18

[81] B. Andersson. *The Lund Model*. Cambridge University Press, 1998. 18

[82] B. Andersson et al. **Parton Fragmentation and String Dynamics**. *Phys. Rep.*, **97 (2 and 3)**:31, 1983. 18

[83] J. Gaudean. PhD thesis, Universiteit Antwerpen, 1984. 19

[84] F. Abe et al. **Measurement of double parton scattering in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV**. *Phys. Rev. Lett.*, **79**:584, 1997. 19

[85] V.A. Abramowsky, V.N. Gribov, and O.V. Kancheli. **AGK Cutting Rules**. *Sov. J. Nucl. Phys.*, **18**:308, 1974. 21

[86] Fritz W. Bopp, R. Engel, and J. Ranft. **Rapidity gaps and the PHOJET Monte Carlo**. 1998. 21, 78, 79, 108

[87] H. Kowalski. **Multiple Interactions in DIS from AGK rules**, 2005. 21

[88] J. Bartels. **Some Remarks on AGK rules**. Mini Workshops on Multiple Interactions and Underlying Events, 2007. 21

[89] C.M. Buttar et al. **The Underlying Event**. 2008. 21

[90] S. Alekhin, G. Altarelli, N. Amapane, J. Andersen, V. Andreev, et al. **HERA and the LHC: A Workshop on the implications of HERA for LHC physics: Proceedings Part A**. 2005. 22

[91] T. Sjostrand and Peter Z. Skands. **Transverse-momentum-ordered showers and interleaved multiple interactions**. *Eur.Phys.J.*, **C39**:129–154, 2005. 22, 23

[92] J. M. Butterworth and M. H. Seymour. **JIMMY4: Multiparton Interactions in HERWIG for the LHC**. 2005. 23, 24

[93] M. Bahr et al. **Herwig++ Physics and Manual**. *Eur. Phys. J.*, **C58**:639, 2008. 23, 26

[94] S. Gieseke, D. Grellscheid, K. Hamilton, A. Papaefstathiou, S. Platzer, et al. **Herwig++ 2.5 Release Note**. 2011. 24, 26

[95] A. Buckley et al. **Systematic event generator tuning for the LHC**. *Eur. Phys. J.*, **C65**:331, 2010. 25, 78

[96] Peter Skands. **Tuning Monte Carlo Generators: The Perugia Tunes**. *Phys. Rev.*, **D82**:074018, 2010. 25, 26, 78

[97] Georges Aad et al. **Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC**. *New J. Phys.*, **13**:053033, 2011. 25

[98] Serguei Chatrchyan et al. **Measurement of the Underlying Event Activity at the LHC with sqrt(s)= 7 TeV and Comparison with sqrt(s) = 0.9 TeV**. 2011. 25

[99] The CMS Collaboration. **Charged particle multiplicities in pp interactions at sqrt(s) = 0.9, 2.36, and 7.0 TeV**. CMS Physics Analysis Summary QCD-10-004, 2010. 25

[100] Rick Field. **Early LHC Underlying Event Data - Findings and Surprises**. *Invited talk at the Hadron Collider Physics Symposium 2010, Toronto, Canada*, 2010. 25, 78

[101] The CTEQ Collaboration. **Global QCD Analysis of Parton Structure of the Nucleon: CTEQ5 Parton Distributions**. *Eur. Phys. J.*, **C12**:375, 2000. 25

[102] J. Pumplin et al. **New Generation of Parton Distributions with Uncertainties from Global QCD Analysis**. *J. High Energy Phys.*, **07**:012, 2002. 25, 26

[103] P. Skands and D. Wicke. **Non-perturbative QCD Effects and the Top Mass at the Tevatron**. *Eur. Phys. J.*, **C52**:133, 2007. 26, 78

[104] T. Sjöstrand, S. Mrenna, and P. Skands. **A brief introduction to PYTHIA 8.1**. *Computer Physics Communications*, **178**(11):852, 2008. 26, 78

REFERENCES

[105] Simon Plätzer. **Herwig++ predictions for CMS**. Private Communication, 2011. 26

[106] C. Flensburg, G. Gustafson, and L. Lönnblad. **Inclusive and Exclusive observables from dipoles in high energy collisions**. *LU-TP 11-13, CERN-PH-TH-2011-058*, 2011. 26

[107] V. N. Gribov. **A REGGEON DIAGRAM TECHNIQUE**. *Sov. Phys. JETP*, **26**:414–422, 1968. 27

[108] David d'Enterria, Ralph Engel, Tanguy Pierog, Sergey Ostapchenko, and Klaus Werner. **The strong interaction at the collider and cosmic-rays frontiers**. 2011. 27, 28

[109] David d'Enterria, Ralph Engel, Tanguy Pierog, Sergey Ostapchenko, and Klaus Werner. **Constraints from the first LHC data on hadronic event generators for ultra-high energy cosmic-ray physics**. *Astropart.Phys.*, **35**:98–113, 2011. 27

[110] David d'Enterria. **Cosmic-ray hadronic MCs vs. p-p MB data at the LHC**. Talk at LHC UE/MB WG, June 2011. 27

[111] H. J. Drescher et al. **A unified treatment of high-energy interactions**. *J. Phys. G: Nucl. Part. Phys.*, **25**:91, 1999. 27

[112] Loyal Durand and Pi Hong. **QCD and rising cross sections**. *Phys. Rev. Lett.*, **58**(4):303, 1987. 27

[113] K. Werner, F. M. Liu, and T. Pierog. **Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron-gold collisions at the BNL Relativistic Heavy Ion Collider**. *Phys. Rev. C.*, **74**:044902, 2006. 27

[114] S. S. Ostapchenko. **Monte Carlo treatment of hadronic interactions in enhanced Pomeron scheme: I. QGSJET-II model**. *Phys. Rev. D.*, **83**:014018, 2011. 27

[115] N. N. Kalmykov, S. S. Ostapchenko, and A. I. Pavlov. **Quark-gluon string model and EAS simulation problems at ultra-high energies**. *Nucl. Phys. Proc. Suppl.*, **52B**:17, 1997. 27

[116] Eun-Joo Ahn et al. **Cosmic ray event generator Sibyll 2.1**. *Phys. Rev.*, **D80**:094003, 2009. 27

[117] A. Capella. **Dual Parton Model**. *Phys. Rep.*, **236**:225, 1994. 27

[118] Stefano Catani. **Aspects of QCD, from the Tevatron to the LHC**. 2000. 29

[119] W. J. Stirling. **LHC Parton Kinetmatics** [online]. 2011. 28, 30

[120] DESY. *Status of Forward Physics Projects at CMS*, 2007. 30

[121] Niladri Sen. *A Study of Multiple Interactions and Underlying Event Structure at the LHC*. Master's thesis, Universite de la Mediterranee, 2007. 30

[122] Michal Deak. *Transversal momentum of the electroweak gauge boson and forward jets in high energy factorisation at the LHC*. PhD thesis, Universität Hamburg, 2009S. 30

[123] **Meeting on Diffraction and Forward Physics at HERA and the LHC**. Talks, 2007. 30

[124] L. Foa. *Phys. Rep.*, **22**(1), 1975. 32

[125] J. Whitmore. *Phys. Rep.*, **27**:187, 1976. 32

[126] J. Whitmore. *Phys. Rep.*, **10**:273, 1974. 32

[127] Hidezumi Terazawa. **Relation between the Average Multiplicity and Transverse Momentum of Produced Particles in Hadron-Hadron Collisions at Very High Energies**. *J. Phys. Society of Japan*, **58**(1):5, 1989. 32

[128] G. Giacomelli and M. Jacob. *Phys. Rep.*, **55**:1, 1979. 32

[129] L. Evans and P. Bryant. **LHC Machine**. *JINST, 3(S08001)*, 2008. 37

[130] The ATLAS Collaboration. **The ATLAS Detector and Physics Performance**. Technical Report CERN-LHCC-99-14, 1999. 37

[131] G.L. Bayatian et al. **CMS physics: Technical design report**. Technical report, CERN, 2006. xi, 37, 38, 47, 65, 74

[132] The ALICE Collaboration. **Technical Proposal for A Large Ion Collider Experiment at the CERN LHC**. Technical Report CERN-LHCC-2001-21, CERN, 2001. 38

[133] The LHCb Collaboration. **LHCb: Technical Proposal**. Technical Report Technical Report, CERN-LHCC-98-004, CERN, 1998. 38

[134] O. Adriani. **LHCf experiment : Technical Design Report**. LHCF-TDR-001 CERN-LHCC-2006-004, 2006. 38

[135] The TOTEM Collaboration. **The TOTEM experiment at the LHC**. *Prog.Theor.Phys.Suppl.*, **187**:281, 2011. 38

[136] Jean-Luc Caron with modifications by Christoph Rosemann. **Layout of the LEP tunnel including future LHC infrastructures**. *LHC Project Illustrations*, 1993. 38

[137] R. Adolphi et al. **The CMS experiment at the CERN LHC**. *JINST*, **03**:S08004, 2008. 40, 45, 48, 51, 52, 53, 54

[138] Dirk Dammann. *Production Cross Section Measurement of Top–Antitop Pairs in the Dimuon Decay Channel at* $\sqrt{s} = 7$ *TeV with the CMS Experiment.* PhD thesis, Universität Hamburg, 2011. 41, 43

[139] G. L. Bayatian et al. **CMS Physics Technical Design Report Volume I : Detector Performance and Software**. *CERN-LHCC-2006-001*, 2006. 44, 46

[140] S. Abdullin et al. **Design, performance, and calibration of CMS forward calorimeter wedges**. *Eur. Phys. J.*, **C53**:139, 2008. 49

[141] V. Andreev et al. **Performance studies of a full-length prototype for the CASTOR forward calorimeter at the CMS experiment**. *Eur. Phys. J.*, **67**:601–615, 2010. 51

[142] R. Pellenq. **Simulations in Physics**. Lectures, 2007. 55

[143] Axel Cholewa. *$D^{*\pm}$ Meson Production at low Q^2 with the H1 Detector Determination of Unintegrated Gluon Densities*. PhD thesis, Universität Hamburg, 2010. 57

[144] John Allison et al. **Geant4 developments and applications**. *IEEE Trans. Nucl. Sci*, **53**:270, 2006. 57

[145] S. Chatrchyan et al. **CMS Physics Technical Design Report, Volume II: Physics Performance**. *J.Phys*, **G34(6)**:995,2376, 2007. xi, 57, 74

[146] S. Agostinelli et al. **GEANT4: A simulation toolkit**. *Nucl. Instrum. Meth.*, **A506**:250–303, 2003. 57

[147] The ALEPH Collaboration. **Performance of the ALEPH detector at LEP**. *Nucl. Instrum. Meth.*, **A360**:481, 1995. 59

[148] The CDF Collaboration. *Phys. Rev.*, **D75**, 2007. [link]. 59

[149] The CMS Collaboration. **Particle Flow Event Reconstruction in CMS and Performance for Jets, Taus and missing tranverse energy**. CMS Physics Analysis Summary PFT-09-001, 2009. 59, 61

[150] J. A. C. Ballin. *Particle Flow at CMS and the ILC*. PhD thesis, Imperial College, London, 2010. 59, 60, 61

[151] The CMS Collaboration. **CMS Tracking Performance Results from Early LHC Operation**. *Eur. Phys. J.*, **C70**, 2010. 59, 103

[152] C. Buttar, J. D'Hondt, M. Kramer, G. Salam, M. Wobisch, et al. **Standard Model Handles and Candles Working Group: Tools and Jets Summary Report**, 2008. * Temporary entry *. 61

[153] J E Huth, N Wainer, K Meier, N J Hadley, F Aversa, Mario Greco, P Chiappetta, J P Guillet, S Ellis, Zoltán Kunszt, and Davison Eugene Soper. **Toward a standardization of jet definitions**. FERMILAB-CONF-90-249-E, 1990. 61

[154] Stephen D. Ellis and Davison E. Soper. **Successive combination jet algorithm for hadron collisions**. *Phys.Rev.*, **D48**:3160–3166, 1993. 61, 62

[155] S. Catani, Yuri L. Dokshitzer, M.H. Seymour, and B.R. Webber. **Longitudinally invariant K_t clustering algorithms for hadron hadron collisions**. *Nucl.Phys.*, **B406**:187–224, 1993. 61, 62

[156] G. P. Salam and G. Soyez. **A practical Seedless Infrared-Safe Cone jet algorithm**. *JHEP*, **0705**:086, 2007. 61, 62

[157] Gavin P. Salam. **Towards Jetography**. *Eur.Phys.J.*, **C67**:637–686, 2010. 62

[158] CMS Collaboration. **Performance of Jet Algorithms in CMS**. Physics Analysis Summary JME-07-003, 2007. 62

[159] M. Cacciari. **Recent Progress in Jets**. Talk at DESY Hamburg, 2009. 62

[160] M. Cacciari and G. P. Salam. **Dispelling the $N3$ myth for the Kt jet-finder**. *Phys. Lett.*, **B641**:57, 2006. 62

[161] M. Cacciari, G. P. Salam, and Soyez. G. **The anti-k_t jet clustering algorithm**. *JHEP*, **0804**:063, 2008. 62

[162] The CMS Collaboration. **Deterimination of the Jet Energy Scale in CMS with pp Collisions at 7 TeV**. CMS Physics Analysis Summary JME-10-010, 2010. 63

[163] K. Kousouris. **Relative Response from Dijet Balance with 7 TeV data**. JEC Meeting, 2010. 64

[164] Hannes Jung. **News**. Internal Meeting, 2011. 66

[165] CMS Collaboration. **Zero bias and HF-based minimum bias triggering for** *pp collisions at 14 TeV in CMS*. CMS Physics Analysis Summary QCD-07-001, 2007. 65

[166] A. J. Bell. **The design and construction of the beam scintillation counter for CMS**. *CERN-THESIS-2009-062*, 2008. 65, 82

[167] T. Aumeyr. **Beam Phase and Intensity Monitoring for the Compact Muon Solenoid Experiment**. Master's thesis, Vienna University of Technology, Austria, 2009. 65

[168] CMS Collaboration. **Tracking and Vertexing Results from First Collisions**. CMS Physics Analysis Summary TRK-10-001, 2010. 67

[169] N. Cartiglia, M. Marone, and R. Arcidiacono. **Measurement of the inelastic pp cross section at $\sqrt{s} = 7$ TeV with the CMS detector using pile-up events**. CMS Analysis Note AN-2011-104, 2011. 67

[170] CMS Collaboration. **Tracking and Primary Vertex Results in First 7 TeV Collisions**. CMS Physics Analysis Summary TRK-10-005, 2010. 67, 103

[171] N. Sen and P. Katsas. **Measurement of the Energy Flow at Large Pseudorapidies for 0.9 TeV and 7 TeV at the LHC**. CMS Analysis Note AN-10-011, 2010. 68, 73, 75, 76, 78, 79, 85, 87, 89, 90, 91, 93, 94, 95, 96, 97, 98, 102, 103, 104, 107, 109, 110, 112, 115

[172] Panos Katsas and Niladri Sen. **Collaborative**. CMS Analysis Note AN-10-011. 68, 72, 73, 76, 89, 90, 91, 93, 94, 95, 96, 97, 102, 104, 107, 112

[173] The CMS Collaboration. **News** [online]. 2010. 69

[174] N. Saoulidou. **Particle Flow Jet Identification Criteria**. CMS Analysis Note AN-10-003, 2010. xi, 70, 71

[175] N. Saoulidou. **PF Jet ID Recommendation for 7 TeV Analysis**. Internal Meeting, 2010. xi, 71

[176] HCAL DPG. **Discussions on HF performance**. Private Communication, 2010. 74, 112, 113

REFERENCES

[177] Vladimir Gavrilov on behalf of the HCAL Detector Performance Group. **Correction Values for Geometrical Non-uniformity**. Private Communication. xi, 74

[178] Alexander Flossdorf and Niladri Sen. **Collaborative**. CMS Analysis Note AN-10-011. 75, 76, 102

[179] A. Proskuryakov. **Observation of diffraction in proton-proton collisions at 900 and 2360 GeV centre-of-mass energies at the LHCParticle Flow Jet Identification Criteria.** CMS AN-10-062, 2010. 77, 103, 104

[180] Frank Cheblana et al. **Optimization and Performance of HF PMT Hit Cleaning Algorithms Developed Using pp Collision Data**. CMS Detector Performance Note DN-2010-008, 2010. 77

[181] Albert Knutsson. **Plotting Script and Asthetics**. CMS Analysis Note AN-10-011. 78, 79, 120, 122, 123, 124, 125

[182] Glen Cowan. *Statistical Data Analysis*. Oxford University Press, 1998. 81

[183] The CMS Collaboration. **Measurement of the Energy Flow at Large Pseudorapidity at the LHC at 900, 2360 and 7000 GeV**. CMS Physics Analysis Summary FWD-10-002, 2010. 92, 101, 102, 110, 112

[184] Taylan Yetkin et al. **Measurement of the Energy Flow at Large Pseudorapidity at the LHC at 900, 2360 and 7000 GeV**. CMS Analysis Note AN-10-002, 2010. 105, 106, 112

[185] Sunanda Banerjee. **Distribution of Dead Material in front of the HF**. Internal Meeting. 105

[186] HCAL DPG. **Issues with HF simulation and reconstruction**. Private Communication, 2010. 106, 113

[187] Pierre Van Mechelen. **Forward QCD studies and prospects at the LHC**. Talk at INT Workshop on Perturbative and Non-Perturbative Aspects of QCD at Collider Energies, September 2010. 108

[188] Dmytro Volyanskyy. **Observation of diffraction with the CMS experiment at the Large Hadron Collider**. 2011. 108

[189] hcalDPG. **Energy scale calibration**. Internal Meeting, 2011. 111

[190] Kerstin Borras, Monika Groethe, and Jeremiah Mans. **Comments from the Analysis Review Committee**. Private Communication. 113

[191] The CMS Collaboration. **Jet Performance in pp Collisions at 7 TeV**. CMS Physics Analysis Summary JME-10-003, 2010. 115

[192] The CMS Collaboration. **Measurement of the energy flow at Large Pseudorapidies for 0.9 TeV and 7 TeV at the LHC**. submitted to JHEP, 2011. * Temporary entry *. 120, 121, 122, 123, 124, 125

[193] Albert Knutsson. **Estimates of hadronisation effects in energy flow**. Private Communication, 2011. 121

[194] Tanguy Pierog, Ralph Engel, Sergey Ostapchenko, and Ralf Ulrich. **Cosmic-ray Monte Carlo generator predictions**. Private Communication, 2011. 121

[195] C. Albajar et al. **A Study of the General Characteristics of $p\bar{p}$ Collisions at $\sqrt{s}$ = 0.2-TeV to 0.9-TeV**. *Nucl. Phys.*, **B335**:261, 1990. 123, 124

[196] Jerome Klapka Jerome. *Three Men on the Bummel*. Serenity Publishers, LLC, 2009. 133

Acknowledgements

The work done as a PhD student and fulfilling the requirements to gain a doctoral degree would not have been possible without the support of many people. I would like to take a moment to mention those who helped me get where I am today. It would be impossible to thank everyone who has supported and encouraged me thus far. Instead of compiling an exhaustive, possibly endless list of names that would be longer than a telephone directory (remember those things...?), I would like to focus on a few good people; the core that have been instrumental in the production of the work presented in this dissertation, both inside and outside the world of physics.

For starters, I would like to thank the people that first got me involved in Particle Physics, tutored me during my initial learning phase and encouraged me to pursue various ventures: Jurgen Brunner, Jose Busto, Pascal Coyle, Christophe Hugon, Laurence Mason, Stephane Poss and Lawrence Sulak. They set the foundations for my work as a PhD student in Particle Physics, and I would not be here if not for their influence during the early period of my university education.

Equally important if not more so is Kerstin Borras. She was the first to get me involved with Particle Physics at DESY, mentoring me through the course of the DESY Summer Student program and an internship during my MSc., which eventually led to my time here as a doctoral student. Without her, many things would not have been possible, and her presence and constant support both as an advisor and a group leader throughout the last few years is greatly appreciated.

For taking me on as his student and integral to the whole experience, I must thank my supervisor, Hannes Jung. I was not the brightest nor the most diligent student, my ideas were often incoherent, and my tendency to leave tasks till the last minute made for a few bumpy moments. His endless patience, continued optimism, infectious enthusiasm for QCD, dedication to his students, and piercing scrutiny of my work is what got me through to the end. At times, I think he read my thesis cover-to-cover more times than I ever did. I am continually surprised (and often inspired) by the

number of hours he consecrates to science. Thanks Hannes, for making it an enjoyable experience and for taking the time to provide counsel at any given moment, regardless of the subject. I would probably not have graduated and published, if not for your wisdom and direction.

I would also like to thank Alexander Flossdorf, Panos Katsas and Albert Knutsson, for their innumerable contributions to this analysis. First, I am very grateful for all the work with which Panos aided me, making the whole investigation a lot less painful. From input for code development, to cross-checking every minute detail, this work would not have been completed (at least not in the given time) if not for his help. Albert and Alex afforded me with a steady stream of practical wisdom, gained from their years in science. Tapping into their knowledge of jets, QCD and event generators saved time and provided an insight I never gleaned from books. More importantly, their help in putting assigned tasks into perspective and translating physics goals into realistic studies bridged the gap between textbook physics and actual research. Not to trivialise their contributions to my life outside DESY, but gratitude for their restaurant recommendations that kept me well-nourished, for regaling me with endless stories and trivia, or for suggesting nutty training regimes and exercises adapted from Spetsnatz should be left for a more appropriate forum.

Though the whole Forward Physics Analysis Group also deserves thanks for this work, I would like to mention Grzegorz Brona specifically for all the effort he put in behind the scenes. As sub-group convenor, he provided a stimulating and fruitful atmosphere to carry out this investigation, lending his extensive knowledge on all areas of CMS to the analysis. His enthusiasm along with his up-to-date knowledge of the detector, physics tools and developments in other studies formed the backbone and primary driving force of the Forward Jets group, and consequently, the measurement presented in this thesis. Of course, I cannot express gratitude towards the Forward PAG without a word of thanks to the CASTOR team; Hans van Haevermaet, Ekaterina Kuznetsova, Pierre van Mechelen, Silvia Ochesanu, Apostolos Panagioutou and Benoit Roland, all of whom helped me get my C-legs. I hope you will excuse the poor pun and weak reference to our cities (Athens, Antwerp and Hamburg) being port towns. A very special thanks to Alan Campbell and Thomas Schroener-Sadenius, also part of the forward groups, for the painstaking task of reading, editing and commenting on the early drafts of my thesis.

I really enjoyed my time here, in large part due to the infrastructure made available to those working at this institution. In particular, I would like to thank the various administrative people at DESY who made my time in Hamburg run so smoothly: Steffi Killough, Bernd Krempl, Nora van Looveren, Ebeling Regina, Catherine Schierholz and Heinrich Zimmermann. By the same token, I would like to express my gratitude to all my colleagues from CMS at DESY and Hamburg University with whom I had the pleasure to work, and all my colleagues at DESY for creating an inspiring, friendly and respectful environment.

Amongst all the colleagues encountered at DESY, special thanks goes to Axel Cholewa, Pedro Cipriano, Michal Deak, Martin von den Driesch, Anastasia Grebenyuk, Federico von Samson-Himmelstjerna, Igor Katkov, Mira Kraemer, Krzysztof Kutak, Lluis Marti-Magro, Zoltan Nagy, Zlatka Staykova, Tobias Toll, Dmytro Volyanskyy - the past and the current participants of our weekly "Physics and Cookies" discussions. From teaching me how to make coffee to making the work place a bit brighter with your distinct personalities, the time shared was much appreciated. I would also like to thank Danny Bot, Pedro Cipriano, Alexander Flossdorf, Johannes Hauk and Sven Schubert, all of whom were my office-mates at one point or another. Each provided an environment appropriate for different stages of my life as a grad student, so I thank you all.

My time here in Hamburg was made more colourful, largely due to the original "four-o-clock" coffee crew from ZEUS (Sarah Boutle, Matthew Forrest, Julia Grebenyuk, Tim Namsoo and Homer Wolf) and their aforementioned H1 counterparts. Thanks for helping me settle into the gloriousness that is Hamburg. I am grateful for the time we spent together, and look forward to the moments our paths may cross again. Friends by association, or add-ons to the coffee-crew, such as Christoph Rosemann, David South, Roman Koegler and Voica Radescu, deserve a word of thanks for agreeing to participate in the viva as "Fragestellers", as well as various other meetings that I will file under miscellaneous; though our interactions were brief and infrequent, they were amongst some of the most memorable.

To the numerous, recurrent characters I met at assorted physics gatherings: understandably, I cannot remember and name you all. Nonetheless, a few come to mind that made my stays at summer schools, conferences and at CERN a bit more bearable. Federico Chavez, Michelangelo Giorgi, Dean Horton, Joerg Mechnich, Edmund Widl

and Kenneth Wraight – discussions with them were as much part of the educational experience that is a doctoral program, as the work itself.

To the friends I made through poker (Dirk Dammann, Timon Mehrling, Thorben Vehling, Wolf Behrenhoff), and the core-group that was ever-present, always willing to lend a hand, available at the drop of a hat at all hours, uttering the oft-heard phrase "one last drink" (Sebastian Aderhold, Clemens Liebig, Mihajlo Mudrinic, Philipp Pahl and Felix Schlander): despite your good intentions and best attempts to prolong my student-life, I seem to have graduated... I cherish the breaks shared and thank you for the company you afforded. With many, I had numerous enjoyable discussions (often peppered with bits of physics), on the subjects ranging from morality to video games, from music and television to whisky and women, and everything else in between; a special salute to any and all for enduring my musings and rants on particle physics, the universe and everything else.

I would especially like to thank Christine and Peter Burmeister who provided me with a lovely home for two years, giving me the comfort, security and stability needed to focus on my work and personal life. The location is perfect, but it is the thought put into making the flat self-sufficient, equipping it with more than just the basic amenities is what makes the place so special. It has been my safe haven and castle for these past two years, and I thank you for opening up this home to me.

Finally, thanks to my family, both immediate and extended, past and present - this work is a testament to the support I received from you all. In particular, my parents who helped me go from wanting to be "a bin-ban man" to "a scientist", constantly believing in me, and opening doors for me through their endless sacrifice; I cannot express how grateful I am for the love of my family, and for everything that went into enabling my dreams and visions to come true. This work is written for you.