MEASUREMENT OF THE UNDERLYING EVENT ACTIVITY IN PP COLLISIONS AT THE LHC USING THE LEADING TRACKS

By

MOHAMMED ZAKARIA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2013

This work is dedicated to my dad, for his vision and unlimited support.

ACKNOWLEDGMENTS

I am very thankful for all the help I have received in learning about the physics and the research tools needed to write this dissertation. Thanks to my adviser Rick Field for his trust, encouragement, and support throughout these years. Many thanks go to Ivan Furic, Darin Acosta, Kevin Burkett, Paolo Bartalini, Luca Mucibillo, and Andrea Lucarouni for their support and inspiration in various forms given throughout this journey, I am in deep gratitude to the depth of their knowledge as well to their giving spirit. Many thanks to my life partner, Justyna Dobrowoloska, for her love and support through the challenges of building a reasonable balance between work and other aspects of life. Thanks goes to my family and friends for their faith and encouragement. In particular, I would like to thank Marwan Suheimat, a living proof of a friendship that survives large (4 dimensional, space like) displacement. A special thanks goes to Rashid Hamdan, a dear friend and room-mate for his great company and friendship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MEASUREMENT OF THE UNDERLYING EVENT ACTIVITY IN PP COLLISIONS AT THE LHC USING THE LEADING TRACKS

By

Mohammed Zakaria

December 2013

Chair: Richard Field
Major: Physics

A measurement of the activity accompanying the hard collision is performed in proton-proton collisions at two center of mass energies of $0.9\ \mathrm{TeV}$ and $7\ \mathrm{TeV}$ at CMS. The direction of the track with the highest transverse momentum is used to divide the azimuthal plane to three regions: “Toward”, “Away”, and “Transverse”. The data are corrected to the hadronic level and compared with two Monte-Carlo Generator tunes. The results are compared with other experiments at the LHC and the Tevatron. These results can be used to add further constrains on the Monte-Carlo generators used to simulate these collisions. It will also enable us to give better predictions for hadronic collisions at the higher energies the LHC is planning to reach in the future.

CHAPTER 1
INTRODUCTION TO PARTICLE PHYSICS

1.1 Overview

Particle physics seeks to understand the fundamental elements of our universe, and the forces between them. Throughout the past two centuries, science achieved a good success in this task by studying the components of matter. After establishing that all chemical elements are made of nuclei and electrons, efforts went to studying the components of these nuclei. By the '30s of the last century, two nuclear components were identified: Protons and Neutrons. The next level was unveiled in the '50s and the '60s with the discovery of quarks and gluons as the building blocks of the physical world as we understand it. During these discoveries, we found more particles that are considered fundamental. These particles, known as leptons, do not have an internal structure according to our current observations, as is the case with quarks and gluons.

1.2 The Standard Model

Most of our knowledge about the subatomic world is organized within a framework called the standard model (Figure 1-1). This model assumes that all currently known particles and forces can be traced to building blocks and force carriers. The building blocks consist of 6 different quarks and 6 different leptons, and there are 4 force carriers (ignoring gravitons and color charge) that facilitate the interaction between these building blocks. The standard model assumes that all fundamental interactions are invariant under local gauge transformations. That is, the Lagrangian for these interactions is invariant under transformations corresponding to conserved quantities. The standard model shows similarities between the different force carriers (photons, Z and W bosons, and gluons), the standard model also presents a program to unify these force carriers in one grand theory, where there is one larger gauge symmetry group that accounts for the known forces. Consequently, force unification assumes that the coupling constants for these forces will become equal at high enough energies. The most recent major

step in this program was the establishment of the electro-weak unification. The current understanding of the electro-weak symmetry requires the existence of a spin 1 boson known as the Higgs boson. The discovery of this particle was one of the main goals of particle physics for the last 4 decades. The expected mass of this particle, to be found after a series of experiments to be more than $100\ \mathrm{GeV}$, made it necessary to build particle accelerators on an unprecedented scale. The Large Hadron Collider (LHC) was designed and built with the purpose of scanning the TeV energy domain to discover the Higgs boson. This target was reached in the summer of 2012 [1], proving the validity of the standard model and the force unification approach.

1.3 The Role of QCD

Most of the collisions at the LHC are dominated by the strong force, which arises due to interactions between the quarks and the gluons (collectively known as partons, a historic term coined by R. Feynman that is used to describe the structerless constituents of hadrons [2]) of the colliding hadrons. It was found from lepton-nucleon scattering experiments that the proton, rather than being a point object, is an extended object filled with many point-like particles. By varying the leptonic probe and the target nucleon, it was found that three of these point particles have spin 1/2, these are known as the valence quarks. It was also found that there are many point particles that exist for a short period of time as flavor neutral pairs known as the sea quarks. The point structure of these partons was deduced from the fact that the form factors associated with them depends only on dimensionless quantities, and not on physical quantities such as Q^2, a property known as the Bjorken scaling.

The partonic model can be further refined by accounting for the fact that quarks and gluons need to have 3 extra degrees of freedom to account to Pauli's exclusion principle in the Δ^{++} particle, where 3 lepton have the same quantum numbers. It should also be taken under consideration that quarks and gluons can emit gluons before or after the

interactions with the probing lepton. This will affect the Bjorken scaling and the shape of the resulting spectrum.

Adding the 3 extra degrees of freedom (referred to as colors) brings us to the theory of Quantum Chromodynamics (QCD). The theory of QCD has a number of similarities to Quantum Electrodynamics (QED). For example, electrons carry the QED charge, and quarks carry the QCD charge, known as color charge. However, there are 3 kinds of color charge versus only one kind of electric charge. There is only one force carrier in QED; i.e. the photon, which has no electric charge attached to it. In QCD, gluons carry the color charge and they are not color neutral. Another difference between QED and QCD is in the coupling constant for each theory. The lagrangian of QCD is given by:

$$\mathrm{L} = \sum_{\mathrm{q}} \bar{\psi}_{\mathrm{q,a}}(\mathrm{i}\gamma^{\mu}\partial_{\mu}\delta_{\mathrm{ab}} - \mathrm{g_s}\gamma^{\mu}\mathrm{t}^{\mathrm{C}}_{\mathrm{ab}}\mathrm{A}^{\mathrm{C}}_{\mu} - \mathrm{m_q}\delta_{\mathrm{ab}})\psi_{\mathrm{q,b}} - \frac{1}{4}\mathrm{F}^{\mathrm{A}}_{\mu\nu}\mathrm{F}^{\mathrm{A}\mu\nu} \ , \tag{1–1}$$

where $\psi_{\mathrm{q,a}}$ are quark-field spinors for a quark with flavor q, mass $\mathrm{m_q}$ and color index a. The term $\mathrm{A}^{\mathrm{C}}_{\mu}$ corresponds to the gluon field and C runs from 1 to 8 to cover the 8 gluons. Term $\mathrm{t}^{\mathrm{C}}_{\mathrm{ab}}$ represents the 8 matrices that are the generators for the $\mathrm{SU(3)}$ QCD group. $\mathrm{g_s}$ is the coupling constant ($\alpha_{\mathrm{s}} = \frac{\mathrm{g_s^2}}{4\pi}$) and $\mathrm{F}^{\mathrm{A}}_{\mu\nu}$ is the field tensor.

In QCD, the coupling constant (α_{s}) is a function of a renormalization scale μ_{R}:

$$\mu^2_{\mathrm{R}}\frac{\mathrm{d}\alpha_{\mathrm{s}}}{\mathrm{d}\mu^2_{\mathrm{R}}} = \beta(\alpha_{\mathrm{s}}) = -(\mathrm{b}_0\alpha^2_{\mathrm{s}} + \mathrm{b}_1\alpha^3_{\mathrm{s}} + ...) \ , \tag{1–2}$$

where β is negative to account for asymptotic freedom. The coefficients $\mathrm{b_i}$ are given for the coupling of an effective theory that counts only quarks much lighter than μ_{R}. Ignoring but the first term, we get a solution for α_{s}:

$$\alpha_{\mathrm{s}}(\mu^2_{\mathrm{R}}) = (\mathrm{b}_0\ln(\mu^2_{\mathrm{R}}/\Lambda^2))^{-1} \ . \tag{1–3}$$

The coupling constant approaches zero at high momentum scale (leading to asymptotic freedom, where the partons act as if they were free particles at high $\mathrm{p_T}$, contrary to QED), and it blows up at small scales (which explains why no single partons can observed), and in between it evolves fast with scale [3], see Figure 1-2. Similar scale dependence is characteristic of many QCD related observables such as parton distribution functions, to be discussed briefly in the Chapter 2.

Many of QCD processes can be calculated using techniques used in QED, provided that the coupling constant is replaced with the strong coupling and that different color combinations are accounted for. This perturbative approach is valid only for large Q^2, which implies a small value for α_s. For large values of α_s the confinement is not well understood.

It is very important to have a good understanding of QCD, and to be able to reproduce QCD effects accurately in our models in order to carry out any analysis at the LHC, as QCD is one of the main sources of background. Achieving this level of understanding requires us to study and model the part of the collisions that does not participate in the hard scattering, which is known as the underlying event. This component of the event affects many of the objects reconstructed by the detector such as jet energies and transverse momenta as well as isolation criteria, especially for the tracking systems.

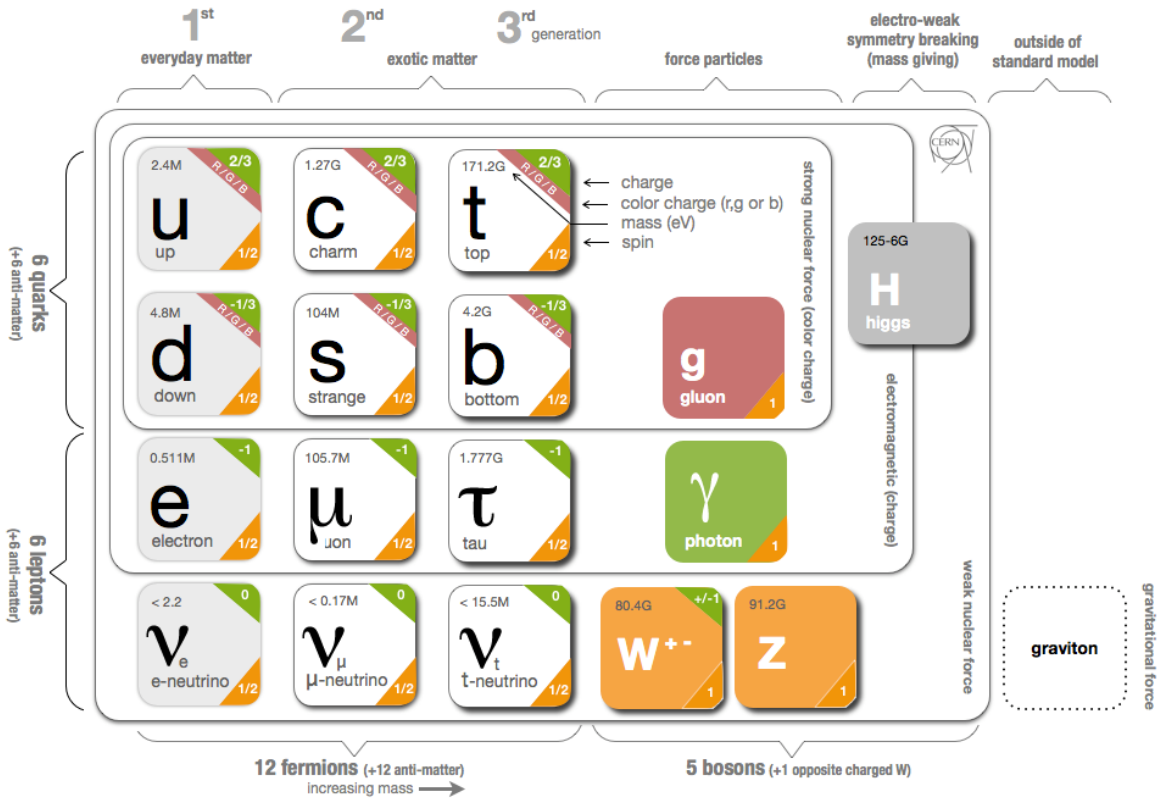

Figure 1-1. Summary of known fundamental particles according to the standard model [4].

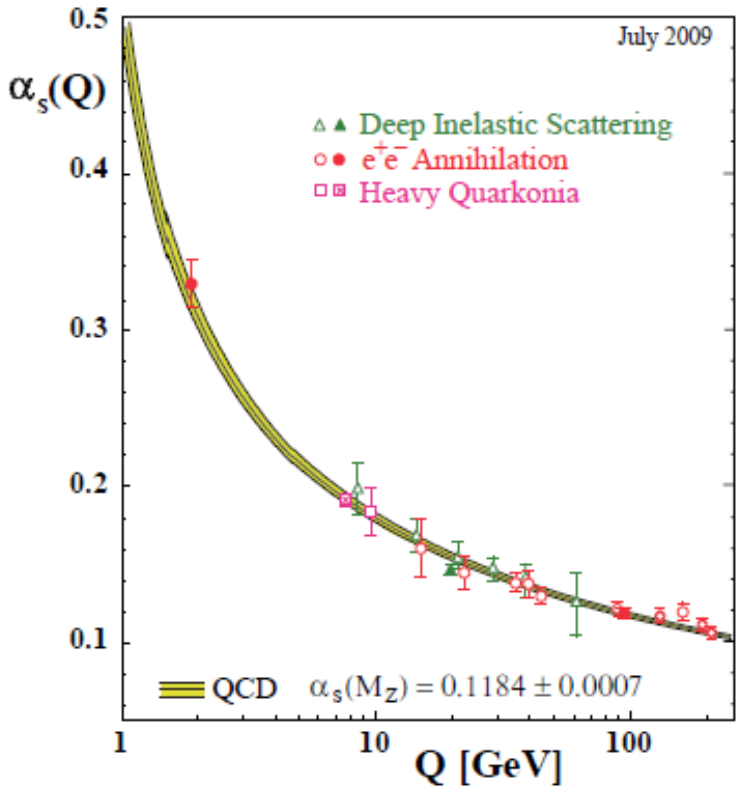

Figure 1-2. The QCD coupling at different scales Q [5].

CHAPTER 2
HADRONIC COLLISIONS

2.1 Overview

In Chapter 2, a brief discussion of various aspects of hadron hadron collisions is pesented. First we discuss the types of hadronic interactions, and the computational approaches used to analyze such complicated processes. In particular, one of the MC generators used to simulate hadronic collisions will be discussed, and the role the UE event plays in this picture. We will conclude with a discussion on tuning MC generators.

2.2 Types of Hadronic Interactions

Proton proton collisions have two major classes: Elastic collisions, where the two interacting protons leave the interaction point as protons after exchanging photons. The second class is Inelastic collisions, where the nature of at least one of the two protons changes due to the exchange of force carriers. Inelastic collisions are the main focus of the experimental investigation at the LHC.

Inelastic collisions can be divided into three groups:

- Single Diffractive (SD), where only one of the two protons loses its structure due to the exchange of a pomeron (i.e. a color singlet exchange).
- Double Diffractive (DD), where both protons lose their structure through the exhange of pomerons.
- Non Diffractive (ND), where both protons interact through weak and strong interactions, and lose their structure.

In this work we will focus only on studying ND events. The values of the cross sections are listed in Table 2-1 [1] .

[1] Another possibility is called central diffraction, where the output is two protons with extra particles. It has a much smaller cross section that it can be ignored in this discussion.

Table 2-1. The approximate cross section for various proton proton interactions. The values are listed at TeV collisions [6]

Interaction type	cross section value [mb]
SD	14.9
DD	9
ND	49.9

The main event in a non-diffractive collision is a 2 $\rightarrow$ 2 scattering between one parton from each hadron. This happens within the size of the hadron and can be calculated using perturbative QCD. See Table 1 in reference [7] for the results. In addition to this process, there are other partonic interactions that are collectively called the underlying event.

The hard component of the event can be calculated using the factorization method [8], where we can relate the interaction at the parton level with the hadronic activity seen at the detector.

$$\sigma(\mathrm{AB} \rightarrow \mathrm{cd}) = \sum_{\mathrm{abcd}} \int \mathrm{dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \hat{\sigma}_{abcd}(\alpha_s^2, \mu_R^2)} \ , \tag{2–1}$$

where a and b run over all quarks, anti-quarks and gluons. $\mathrm{f_{a/A}}$ ($\mathrm{f_{b/B}}$) are the parton density functions that give the probability of having parton 'a' carrying fraction $\mathrm{x_a}$ ($\mathrm{x_b}$) of the proton momentum at scale μ_F^2 known as the factorization scale. It serves to bridge between the perturbative quantities such as $\hat{\sigma}$ and non perturbative quantities such as PDFs [9] and μ_R^2, which is known as the normalization scale.

2.3 Parton Distribution Functions

Parton distribution functions (PDF) are a result of hadrons being composite objects. They can be extracted from structure functions, a measurable physics quantity, within the limits of a specific factorization scheme i.e: Order by order in the perturbation expansion. At the leading order, PDFs have the simple interpretation of being the partonic distribution of the hadron in momentum space. They are related to the structure function $\mathrm{F_2}$ using the following formula:

$$F_2(x, Q^2) = \sum_i e_i^2 x f_i(x, Q^2) \ . \tag{2–2}$$

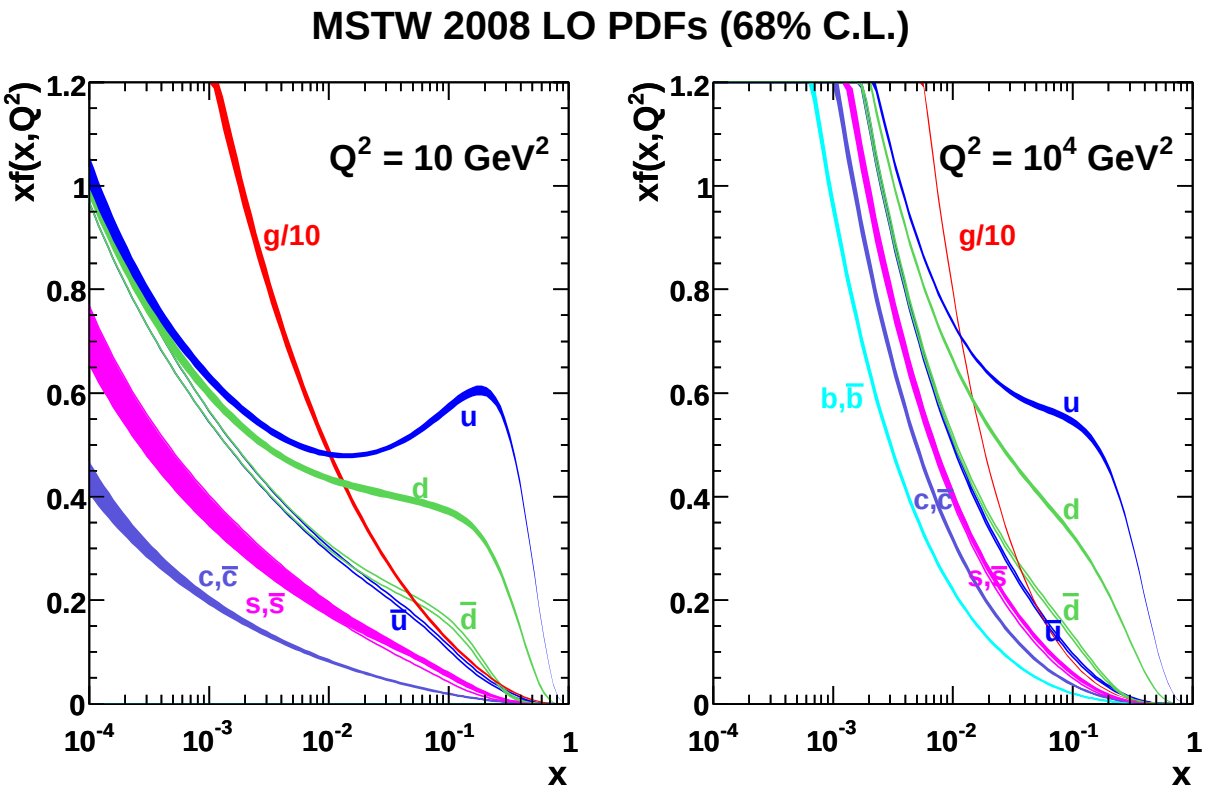

Figure 2-1. PDF to LO for the proton according to MSTW model at two energy scales [10].

Many aspects of PDFs are still under study, such as PDF behavior at low x regions, PDF uncertainties, and how to use them in conjunction with parton-shower MC codes, see reference [9] for a more detailed discussion. There are many models available for PDFs based on the methods used to parametrize the data. One example is shown in Figure 2-1.

Determining the proton's PDFs involves data from various experiments, some with direct sensitivity to quarks such as Deep Inelastic Scattering (DIS), the other with

indirect sensitivity to gluons such as DIS energy evolution, and other experiments with direct sensitivity to both quarks and gluons such as jet data [3].

2.4 QCD Calculations Using MC Generators

There are two general methods to evaluate Equation (2–1): The first one relies on performing the calculations for a fixed number of perturbative orders for the partonic cross section, the second one uses QCD MC models to generate partonic showers.

2.4.1 The Fixed Order Method

Leading Order (LO) MC generators such as ALPGEN [11] and MADGRAPH [12] use tree level diagrams, then carry the integration over the desired phase space and calculate the squared matrix elements for each partonic subprocess. This level of calculations can account reasonably for $2 \rightarrow$ 6-8 processes (two particles interacting to produce up to 6 to 8 particles).

The next level of calculations, known as the Next Leading Order (NLO) level adds one loop to the Feynman diagrams used in the calculations, a step that adds the additional task of regularization (i.e. getting rid of the infinities associated with the emission of soft gluons by integrating the loop momenta over 4 - ϵ dimensions rather than 4). After that there is a need to account for the fact that the physical cuts introduced in a real experiment are in 4 dimensions. The standard method is to introduce suitable counter terms that do not affect infrared (IR) safe observables. These terms will cancel all collinear and soft divergences and can be integrated analytically in the 4 - ϵ space. NLO order calculations are very lengthy and consume large CPU resources. Most of the available models cover $2 \rightarrow$ 3-5 processes.

The Next to Next Leading Order (NNLO) level suffers the same problem that NLO calculations have. But this time we are dealing with two soft and two collinear divergences in the diagrams. The subtraction procedure is thus more complicated and we have routines that only deal with $2 \rightarrow 1$ level calculations for hadronic interactions.

2.4.2 The Parton Showering Method

The second method depends on trying to re-interpret the divergences observed in the first approach (IR and collinear divergences) physically. The first step is to calculate the probability of a parton to radiate a gluon above some transverse scale. To a LO this can be written as [3]:

$$\mathrm{P}(\text{emission above } \mathrm{k_T}) \simeq -\frac{2\alpha_\mathrm{s}\mathrm{C_F}}{\pi}\int\frac{\mathrm{dE}}{\mathrm{E}}\int\frac{\mathrm{d}\theta}{\theta}\Theta(\mathrm{E}\theta-\mathrm{k_T}) \ , \tag{2–3}$$

Where P is the probability of emitting a gluon with a transverse momentum above $\mathrm{k_T}$, α_s is the strong coupling constant and is related to the probability of a QCD emitting/absorbing a gluon, and $\mathrm{C_R}$ is the color factor, which is directly related to the strength of the coupling between two different color states through a single gluon exchange. Subtracting the previous result from one gives us the probability of not having a gluon emission. At the soft and the collinear limit, this result can be expanded to all orders:

$$\mathrm{P}(\text{no emission above } \mathrm{k_T}) \equiv \Delta(\mathrm{k_T},\mathrm{Q}) \simeq \exp[-\frac{2\alpha_\mathrm{s}\mathrm{C_F}}{\pi}\int\frac{\mathrm{dE}}{\mathrm{E}}\int\frac{\mathrm{d}\theta}{\theta}\Theta(\mathrm{E}\theta-\mathrm{k_T})] \ , \tag{2–4}$$

the term $\Delta(\mathrm{k_T},\mathrm{Q})$ is a simplified version of the Sudakov factor [13], were we assumed α_s to be constant and took it out of the integral, and used a hard collinear radiation term $\mathrm{dE/E}$ instead of a full collinear splitting function. The Sudakov form factor allows us to calculate the distribution of the transverse momentum $\mathrm{k_{T1}}$ of the gluon with the largest transverse momentum in the hadronic event:

$$\frac{\mathrm{dP}}{\mathrm{dk_{T1}}} = \frac{\mathrm{d}}{\mathrm{dk_{T1}}}\Delta(\mathrm{k_T},\mathrm{Q}) \ . \tag{2–5}$$

This leads to the MC part; we can take a random number r from a uniform distribution bound between 0 and 1 to be the Sudakov factor, then solve the Equation (2–5) for $\mathrm{k_{T1}}$. This gives us a gluon in addition to the initial state of the system. We repeat the procedure to generate a gluon with $\mathrm{k_{T2}} < \mathrm{k_{T1}}$. The procedure is repeated until we reach some cut-off scale, and thus the parton shower has been generated.

The previous discussion covered what is known as $\mathrm{p_T}$ ordered showers. This is the procedure used in many known MC generators such as PYTHIA 8 [14] and PYTHIA 6.4 [15] as well as SHERPA 1.2 [16]. Other ordering variables are used in generators such as older versions of PYTHIA 6, which are based on virtuality ordered showers, and angular ordering that is used in another class of generators such as HERWIG [17]. Parton showers are also accompany incoming partons, and these need a more careful treatment to account for PDFs.

The next step in the parton shower method is hadronization, which is also a non-preturbative process that is model based in MC generators. Two known models are the Lund string model used in PYTHIA and the cluster model used in SHERPA and HERWIG. More details are provided in their manuals cited in the previous paragraph.

2.5 The Underlying Event

The term Underlying Event (UE) denotes to any additional activity beyond the basic hard process in the event and its associated initial state and final state radiation (Denoted by ISR and FSR respectively). The dominant contribution to the UE is believed to come from additional color exchange between the colliding partons. In particular, through multiple parton interactions (MPI) or pomeron cuts [18] and beam-beam remnants (BBR). Figure 2-2 is an illustration of the UE in a hadronic collision. It is impossible to experimentally distinguish the activity coming from the various contributions on an event by event basis. Instead, we utilize different methods to study phase space regions more sensitive to some of its components as we shall explain later.

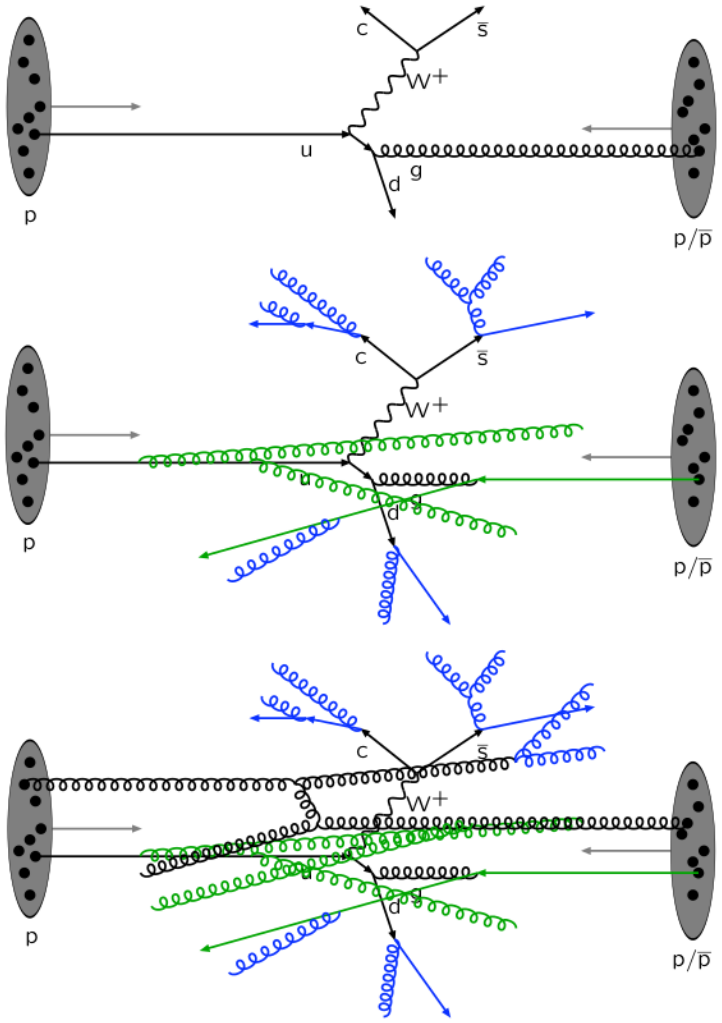

Figure 2-2. A schematic diagram showing the UE in a hadron collision [19]. The upper panel represent the hardest $2-2$ interaction, the second panel shows ISR and FSR in blue and green. The third panel shows the combined effect of the previous two panels as well as additioal partonic interactions.

The addition of MPI leads to the possibility of observing extra parton-parton interactions. The additional jets are referred to as minijets. These jets have the property of forming back-to-back pairs, compared to jets from bremsstrahlung that are aligned with the direction of the parent partons. This extra activity is more observable at low $\mathrm{p_T}$ and gives significant contribution to the color flow and scattering energy of the event. This is observed in the detector as an increase in multiplicity and $\Sigma\mathrm{E_T}$, and as a contribution to the break-up of the beam remnants in the forward direction.

The first MC model for MPI was proposed in the late 80's [20]. The main premise behind it was to view the interacting hadrons as a beam of incoming partons. The scenario of more than one interacting parton can then be treated probabilistically. This interaction will also change the color flow of the whole system.

Quantitatively, having MPI that are mostly soft in scale means that the $\hat{\mathrm{t}}$-channel is almost on shell and thus the partonic cross section behaves roughly as:

$$\mathrm{d}\hat{\sigma} \propto \frac{\mathrm{d}\hat{\mathrm{t}}}{\hat{\mathrm{t}}^2} \sim \frac{\mathrm{d}\hat{\mathrm{p}}_\mathrm{T}^2}{\hat{\mathrm{p}}_\mathrm{T}^4} \ . \tag{2–6}$$

We integrate the cross section from a cut-off value to the center of mass energy using the leading order 2 $\rightarrow$ 2 matrix with PDFs included. The result is shown in Figure 2-4 and it shows the cross section for this simple MPI model exceeding the total cross section for the regions of $\mathrm{p_T}$ below 5 GeV/c [21]. This can be explained by recalling the fact that while the total interaction cross section counts an event as one, the 2 $\rightarrow$ 2 cross section might count it more than once due to MPI. Hence, we can relate the two cross sections:

$$\sigma_{2\rightarrow 2(\mathrm{p_{Tmin}})} = \langle \mathrm{n} \rangle (\mathrm{p_{Tmin}}) \sigma_{\mathrm{tot}} \ , \tag{2–7}$$

and the probability of having number n of interactions in one event is given by:

$$\wp_n(p_{Tmin}) = [\langle n \rangle (p_{Tmin})]^n \frac{\exp[-\langle n \rangle](p_{Tmin})}{n!} \quad . \tag{2–8}$$

This mathematical approach can be further refined by imposing momentum conservation, so that partons will not use more momentum than what the parent hadron had. This serves to suppress the large divergence as p_T goes to zero. One also has to consider color screening. That is, at low p_T the wavelength of the exchanged parton becomes larger than the typical color-anticolor separation distance, and its resolving power is reduced to seeing the color average charge that vanishes as p_T goes to zero, effectively giving an IR cut-off for the interaction. This scale can be estimated using the following approximation:

$$p_{Tmin} \simeq \frac{\hbar}{r_p} \approx \frac{0.2 \text{ GeV . fm}}{0.7 \text{ fm}} \approx 0.3 \text{ GeV} \approx \Lambda_{QCD} \quad . \tag{2–9}$$

In reality, the value obtained in Equation (2–9) is a bit too small. In MC models this number is effectively set as a parametet that can be inserted to Equation (2–6) as a step function (as is done in HERWIG) or as a smooth function as in Equation (2–10) to regularize the divergence further (as is the case in SHERPA and PYTHIA). This parameter can be energy dependent and is one of the key parameters to consider when exploring new energy regions. Higher energies mean that PDFs can be explored at smaller x values (as shown in Figure 2-1), where the number of partons drastically increases and they become more packed; i.e. smaller color screening distance. This leads to a major concern regarding the uncertainties in the energy and the x scaling of the cut-off when extrapolating between different collider energies [9].

One of the main features of the UE is the so-called ‘pedestal effect’; the observation that hard interactions appear to be accompanied with much higher UE activity than events with no hard interactions. This is interpreted in terms of the impact parameter

of the colliding hadrons [22], where having a hard scattering means that the impact parameter is small enough that more partons are engaged at a smaller x as well. Thus the ability to describe both types of collisions (called central and peripheral collisions, see Figure 2-3) depends on the quality of the models of the impact parameter dependence. Special attention is given when interpreting the results of the zero bin of the Poisson distribution, which corresponds to an impact parameter larger than the diameter of the hadron (no ND interactions). For MC models that describe the entire inelastic and non-diffractive cross-section, this bin is simply ignored as it represents diffractive or elastic scattering that are modeled separately. For MC models that are restricted to hard inelastic events, this can be reinterpreted as the fraction of the total inelastic cross-section that has no hard interactions.

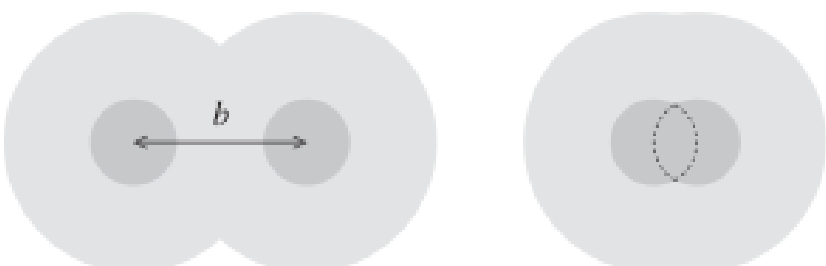

Figure 2-3. A schematic representation for two protons passing by vertically in a peripheral collision, where the impact parameter is large enough that only the soft part of the protons is involved in the collision (left) and two protons involved in a head on collision, where the impact parameter is mall that the core of the protons participates in the collision (right) [22].

$$\alpha_{\mathrm{s}}(\mathrm{p}_{\mathrm{T}}^2)\frac{\mathrm{dp}_{\mathrm{T}}^2}{\mathrm{p}_{\mathrm{T}}^4} \rightarrow \alpha_{\mathrm{s}}(\mathrm{p}_{\mathrm{T0}}^2+\mathrm{p}_{\mathrm{T}}^2)\frac{\mathrm{dp}_{\mathrm{T}}^2}{(\mathrm{p}_{\mathrm{T0}}^2+\mathrm{p}_{\mathrm{T}}^2)^2} \quad . \tag{2–10}$$

Further enhancements to the MPI model include showering the partons involved in this activity and perturbative rescattering. While the showering of the output follows similar pattern as for $2 \rightarrow 2$ scattering process, the showering of partons before entering the MPI interaction region raises questions on how to correlate multi-parton density,

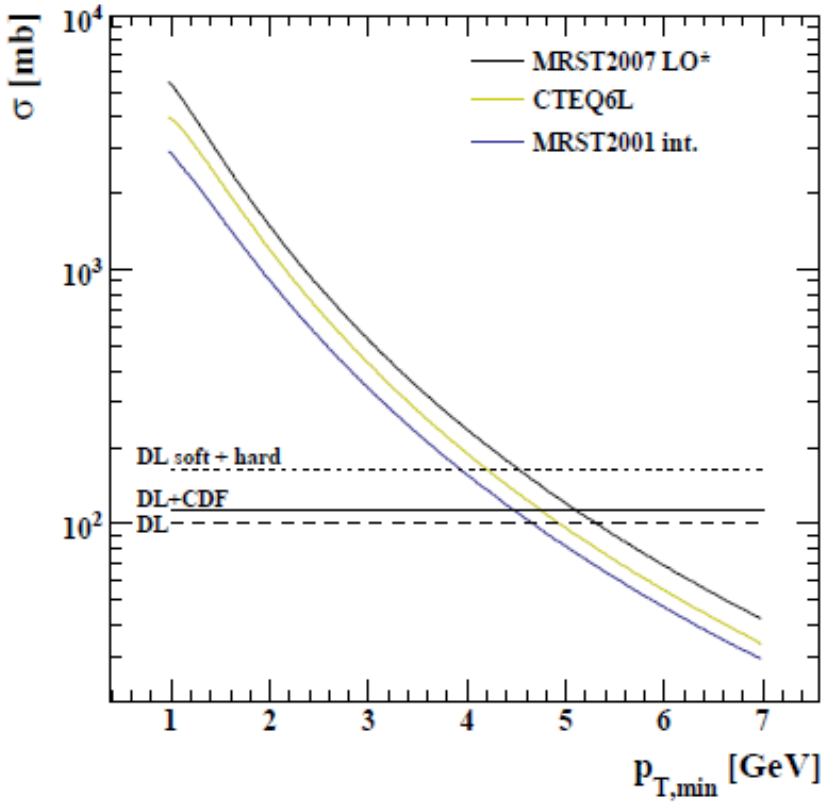

Figure 2-4. The cross section obtained with $2 \rightarrow 2$ leading order calculation vs the total cross section (the horizontal line) using several PDF models [21].

which is a topic of ongoing research. Perturbative rescattering occurs when partons are allowed to do several interactions, which is also an active line of research [23].

Introducing the MPI to the system adds challenges to hadronization models, since it has to color neutralize different color systems (beam remnants, as well as the primary interaction) that are separated in the rapidity space. Many of the IR sensitive variables (hadron multiplicities and hadron spectra, for example) depend crucially on these correlations in color space. To date, there is a large uncertainty on how to address this matter which is reflected in the substantial amount of variation between different MC models.

2.5.1 Studying the UE in MC Generators

The main feature of the parton showering MC models is their ability to provide a complete and detailed picture of the collider final state [24]. The level of accuracy depends on the chosen process (the more inclusive the process is, the better the MC description will be) and the level of detail the simulation contains. This is obvious from the improvement observed within the MC generators as better theoretical models become available; for example, including matching higher order matrix elements and better non-perturbative models. It also depends on constraining the free parameters of these models using data, a process known as generator tuning. As an example, we will discuss some of the tunning aspects of one of the known MC generators: PYTHIA 6.4.

In this class of MC generators, a unified approach is adopted in modeling all inelastic and ND events [24]. That is, the minimum bias (MB) event is considered as a soft limit case of dijet production and the UE events accompanying it, with no change in models between the two. There are many parameters associated with the regularization of the MPI cross section:

- The Infrared Regularization scale: As was explained earlier and in Equation (2–10), the parameter $\mathrm{p_{T_0}}$ sets the scale for the color screening effect and is called the infrared regularization scale. In PYTHIA, it is set to have a power low function of center of mass energy:

$$\mathrm{p_{T0}}(\sqrt{\mathrm{s}}) = \mathrm{PARP}(82).\left(\frac{\sqrt{\mathrm{s}}}{\mathrm{PARP}(89)}\right)^{\mathrm{PARP}(90)} , \qquad (2–11)$$

  where $\mathrm{PARP}(82)$, $\mathrm{PARP}(89)$, $\mathrm{PARP}(90)$ are PYTHIA parameters that can be tuned: $\mathrm{PARP}(82)$ sets the value of $\mathrm{p_{T_0}}$ at the chosen center of mass energy as determined by $\mathrm{PARP}(89)$, $\mathrm{PARP}(90)$ determines how steep the change of $\mathrm{p_{T0}}$ is with as the center of mass energy changes, see Figure 2-5.

- The Transverse Mass Distribution: Another important aspect for tuning is the shape of the proton matter distribution. This stems from the fact that the amount of MPI interactions in a given collision is proportional to the amount of matter overlap between the colliding beam particles. That is, the smaller the impact parameter b is, the more MPI activity is to be expected as is demonstrated by Figure 2-3. If the proton structure is uniform, the difference between peripheral and central collisions will be small. On the other hand, a strongly peaked distribution can make

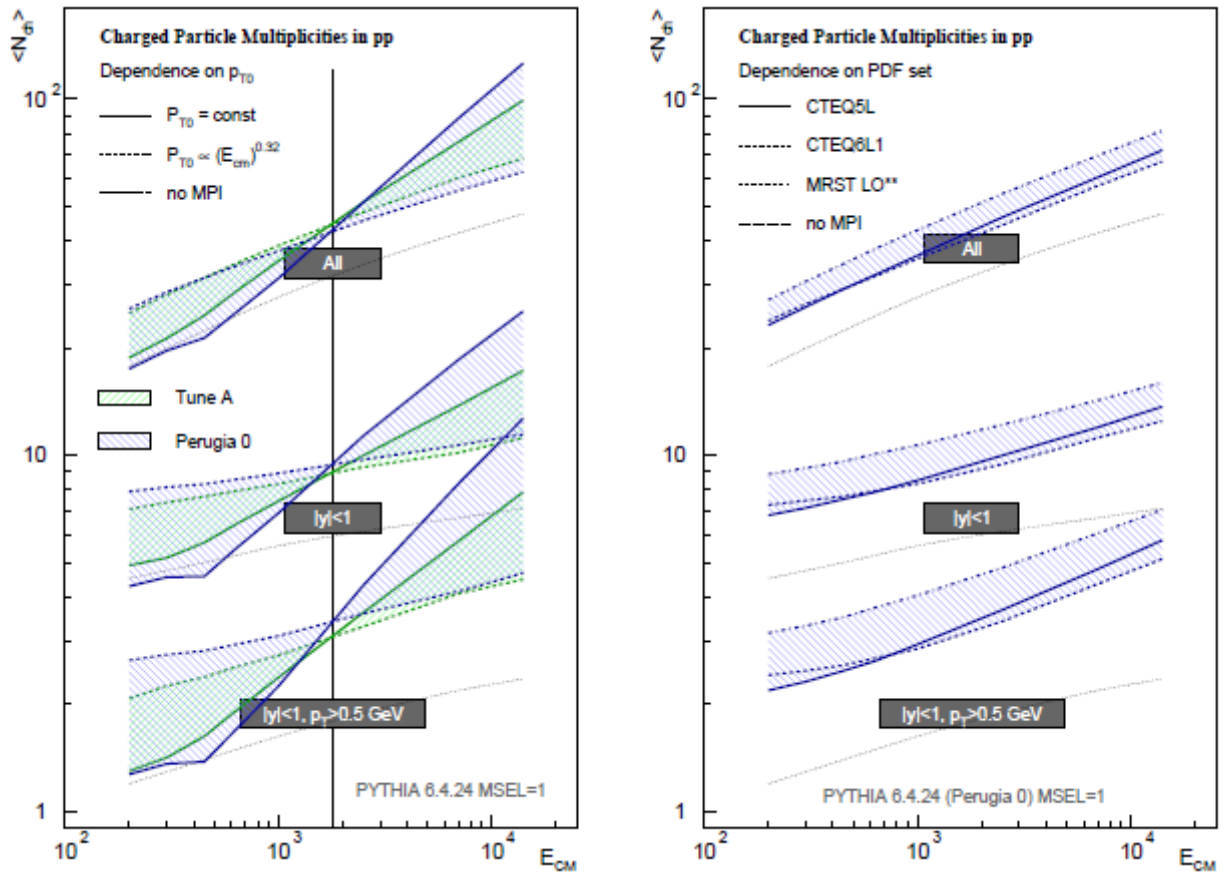

Figure 2-5. The charged particle multiplicity for few phase space regions using different MPI schemes [24].

the activity in central collisions much higher than in peripheral ones. The matter distribution is modeled according to the following:

$$\mathrm{D(b)} \propto \exp(-\mathrm{b^d}) \ , \tag{2–12}$$

where the power d is a free parameter whose range is normally taken to be from $\mathrm{d} = 1$ (exponential, representing a very peaked structure) to d = 2 (Gaussian, representing a smooth structure). Note that the normalization of this distribution is fixed to unity. Note also that b is given in an arbitrary unit; since the only dimensionful quantity is the total cross section, the b shape does not affect the total cross section at all in this type of model, and only the dimensionless ratio $\mathrm{b}/ < \mathrm{b} >$ appears in the explicit calculations. The power, d, appears as the parameter PARP(83) in PYTHIA. It is not assumed to change with energy:

$$\mathrm{d}(\sqrt{\mathrm{s}}) = \mathrm{PARP(83)} \ . \tag{2–13}$$

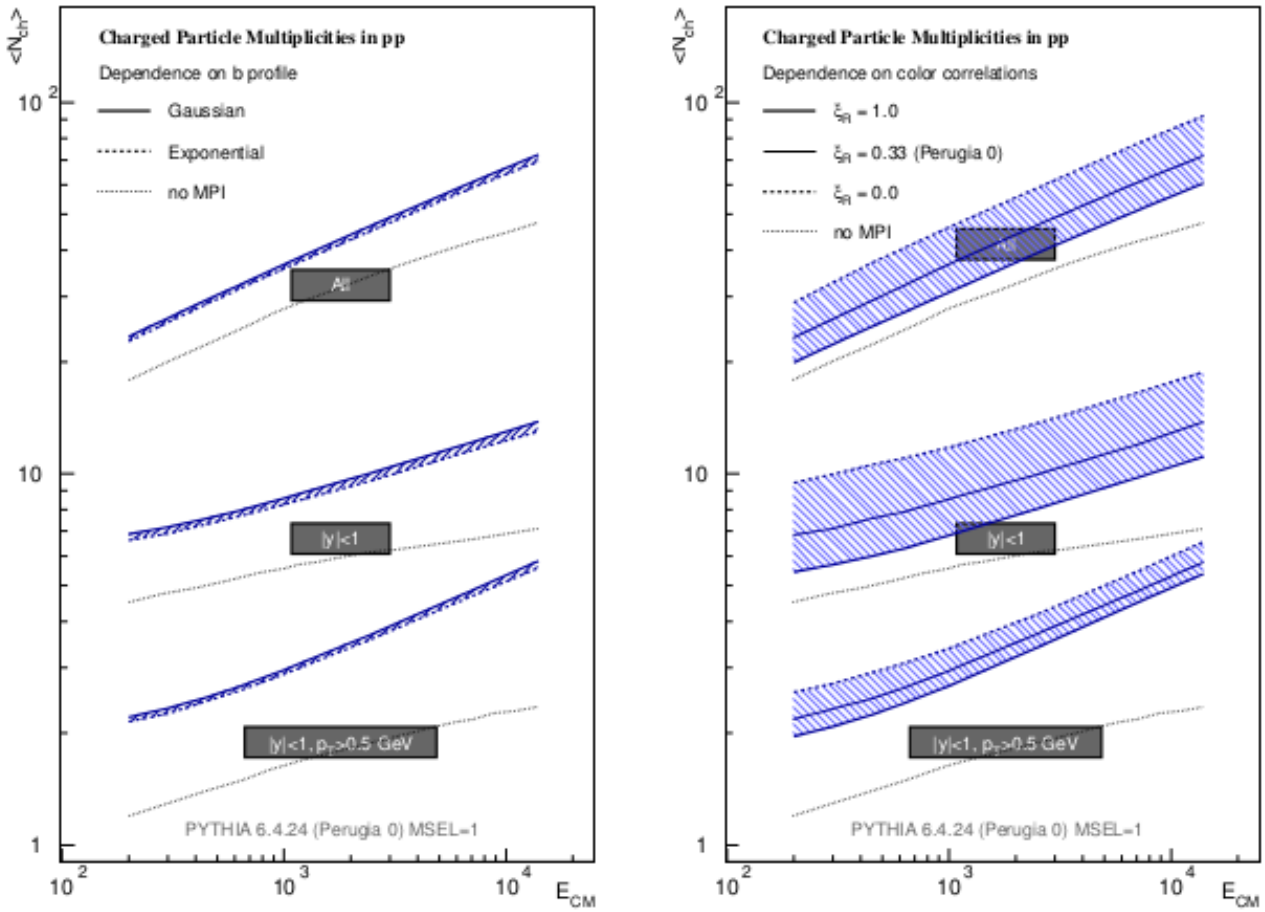

Figure 2-6. The charged particle multiplicity for few phase space regions using different matter distribution and color correlation schemes [24].

In other words, the IR regularization term affects the average activity, and the transverse mass distribution term controls the deviation from this average for peripheral and central collisions [24]. It was found from data that the overlap function is closer to a gaussian [24].

- Color Reconnection Strength: MPI correspond to one or more color exchanges between the partons that gives a color neutral final state at the hidronization phase. In PYTHIA this is done, as mentioned earlier, through the Lund string model. In particular, we assume that a large, color neutral, string that represents the unstable state gets broken into smaller strings that are color neutral as well. Such scenario seems necessary to account for the increase of mean $\mathrm{p_T}$ track multiplicity in minimum bias events [24]. The parametrization of the Lund model is accomplished with the string interaction strength (R), and is set by $\mathrm{PARP(78)}$. The larger the parameter is, the harder the $\mathrm{p_T}$ spectrum will be.

2.5.2 Fundamentals of Tuning

Due to the large number of parameters available in current MC generators, there is a need for a framework to automate the tuning procedure. There have been 3 main approaches toward MC tuning in the last couple of decades:

- Manual tuning: Where an expert eye adjusts few selected parameters of the MC to fit the behavior required. Such manual approach requires a significant insight towards the algorithmic response to parameter choices, which can be developed through experience. It is intrinsically slow as it involves many iterations of parameter choices before getting a reasonable output with satisfying statistics. A typical tuning effort normally comprises of adjusting only few parameters and observing few plots for responsiveness. For each new generator, there is a need to start the learning process from scratch. Despite all that, manual tuning still maintains strong presence among the MC community. In fact, most of the tunes premiered in previous Compact Muon Solenoid (CMS) UE analyses had manual tunes such as tunes $\mathrm{Z1}$ and $\mathrm{D6T}$.

- Brute Force tuning: Which covers any approach aiming at dividing the parameter space to points and running the MC generator at each one of them. The problems with such approach are obvious; selecting only few parameters with, say, 10 points each will require thousands and thousands of MC runs. Even then, the sampling granularity might mask any meaningful results. Another method within this strategy is to rely on the best χ^2 fit. Also hindered by the limited scale we can cover and the lack of ways to improve he best fit, or to distinguish between local and global minima.

- Parametrization-based tuning: This strategy depends on parametrizing the generator behavior, where we fit the polynomial to the generator response of each observable bin to a certain element of the parameter vector. Once this is done, we construct a goodness of fit function and minimize it. The result is a parameter vector that should, in principle, be able to predict the best description of the tunning data the generator can provide. This can be done in a considerably shorter time compared to the previous two approaches.

Tune $\mathrm{Z2^*}$, introduced in our results, used the third approach. This was done using the PROFESSOR [25] tunning framework, which relies on constructing fast analytic models of the generator by analysing its response to shifts in the parameter space. The comparison to experimental data is done via Rivet [26] analysis tool. The tune was based on fitting the profile plots for CMS analysis that used the leading track-jet to study

the UE at CMS [27]. Only two parameters were tuned: PARP(90) and PARP(82), and all profile plots were given equal weight for the two center-of-mass energies.

CHAPTER 3
THE COLLIDER AND THE DETECTOR

3.1 The Large Hadron Collider

The LHC is the world's highest energy collider. The highest energy proton-proton collisions have reached the value of $4\ \mathrm{TeV}$ per beam proton in 2012 and is expected to reach $13\ \mathrm{TeV}$ when it resumes operations in 2015. The running protons in the ring form a beam current of approximately $0.58\ \mathrm{A}$. The two proton beams run in separate pipes except at interaction segments of $120\ \mathrm{m}$ at each interaction point. This would lead to around 30 head-on collisions at each segment.

The collider represents the last link in a system of smaller accelerators. Each accelerator rises the energy of the beam until it reaches the next stage. Table 3-1 lists various accelerator systems used to reach the TeV energy scale at the LHC.

Table 3-1. The different systems used to accelerate hadrons at the LHC.

Kinetic Energy of the proton	Accelerator
$50\ \mathrm{MeV}$	Linac 2
$1.4\ \mathrm{GeV}$	PS Booster
$25\ \mathrm{GeV}$	PS
$450\ \mathrm{GeV}$	SPS
$7\ \mathrm{TeV}$ (planned)	LHC

The LHC ring is made of eight arcs. Each arc contains 154 dipole binding magnets. Connecting the arcs are insertion points that consist of a long straight section with a transition region at each end, called the dispersion suppressors. The exact layout of the straight section depends on the insertion point specific use: Physics, injection, beam dumping, or beam cleaning.

Since the goal of the LHC is to produce processes rarely observed in nature at unpreceded energies, the most important parameters are the beam energy and the number of interesting events. This is expressed through the relation between the number of events and the luminosity:

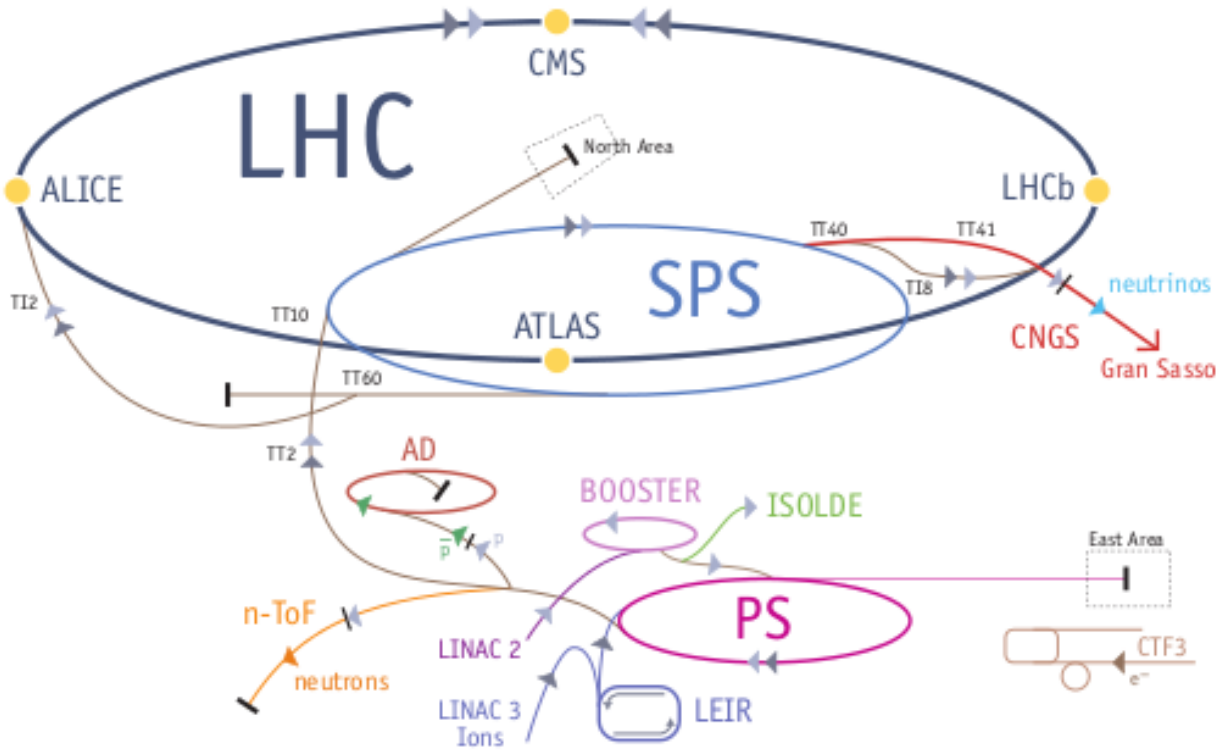

Figure 3-1. A schematic diagram for the LHC chain of acceleration [28].

$$\mathrm{N_{event}} = \mathrm{L}\sigma_{\mathrm{event}} \quad , \tag{3–1}$$

where L is the integrated luminosity and σ is the cross section for the process under investigation. The luminosity depends on the beam parameters and can be written, assuming a Gaussian beam distribution, as [29]:

$$\mathrm{L} = \frac{\mathrm{N_b^2 n_b f_{rev} \gamma_r}}{4\pi\epsilon_{\mathrm{n}}\beta^*}\mathrm{F} \quad , \tag{3–2}$$

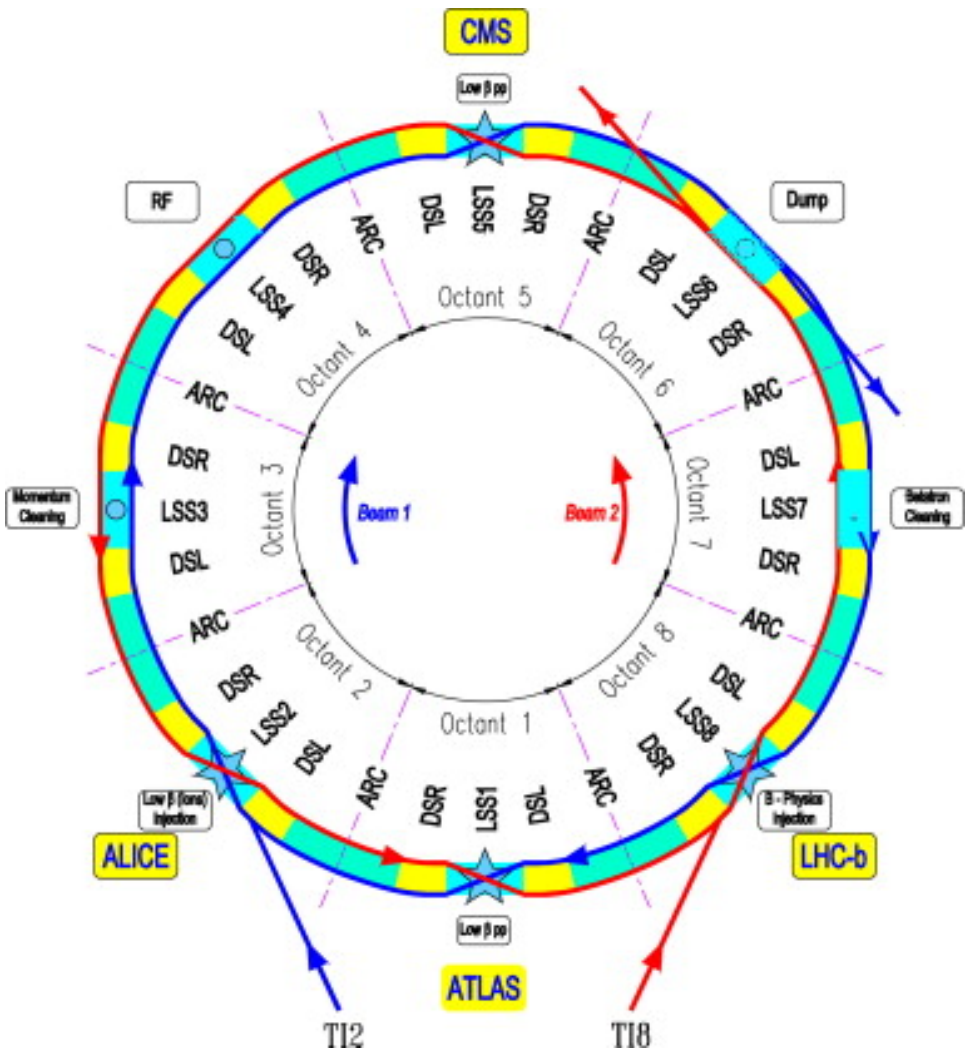

Figure 3-2. A diagram showing the layout of the LHC [30].

where $\mathrm{N_b}$ is the number of hadrons per bunch from one beam, $\mathrm{n_b}$ is the number of bunches in each beam, $\mathrm{f_{rev}}$ the revolution frequency, γ_r is the relativistic gamma factor, ϵ_n is the transverse normalized emittance, β^* is the optical beta function at the interaction point, and F is the geometry luminosity reduction factor [29]:

$$\mathrm{F} = \left(1 + \left(\frac{\theta_\mathrm{c}\sigma_\mathrm{z}}{\sigma^*}\right)^2\right)^{-\frac{1}{2}} \quad , \tag{3–3}$$

where θ_c is the crossing angle at the interaction point, σ_z is the bunch length and $\sigma^* = \sqrt{\beta^*\epsilon_\mathrm{n}}$ is the rms of the beam size at IP. The LHC aims to deliver luminosity of the order of $10^{34}\ \mathrm{cm^{-2}s^{-1}}$ at CMS and ATLAS, with 2808 bunches and bunch spacing of

$25\ \mathrm{ns}$, the maximum acceptable value for ϵ_{n} is $3.75\ \mu\mathrm{m}$ and the beam dynamics as well as the mechanical aperture sets the limit on the number of particles per bunch to be $1.15\mathrm{x}10^{11}$.

3.2 The Compact Muon Solenoid

CMS is one of the major experiments that utilize the LHC beam. Its main feature is the $3.8\ \mathrm{T}$ superconducting magnetic solenoid. The solenoid surrounds the tracking system and the calorimetric systems and immerses them with its uniform magnetic field. The returning yoke contains the muon detection system. Each of the detector systems can be imagined as a cylinder with a barrel parallel to the beam line, plus two caps on the ends. This design ensures almost a full angular coverage, as shown in Figure 3-3. The conventional coordinate system places the point of origin at the interaction point. The x-axis is parallel to the Earth's surface and points toward the LHC's center. The y-axis points upward, and the z-axis is parallel to the beamline. The azimuthal angle (ϕ) is measured in the x-y plane with respect to the x-axis. The polar angle (θ) is measured in the z-y plane with respect to the z-axis. For convenience θ is replaced with pseudo-rapidity (η), which is defined as:

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \quad , \tag{3–4}$$

Below we list the major components of CMS with a brief description of each part:

3.2.1 The Superconducting Magnet

The magnet provides a strong and uniform magnetic field for momentum measurement at CMS. The magnetic coil operates at $4.5\ \mathrm{K}$ using liquid helium. The iron yoke secures a returning field of about $2\ \mathrm{T}$, with a large bending power for accurate momentum resolution for the muons. At full current, $2.6\ \mathrm{GJ}$ of energy is stored in the system.

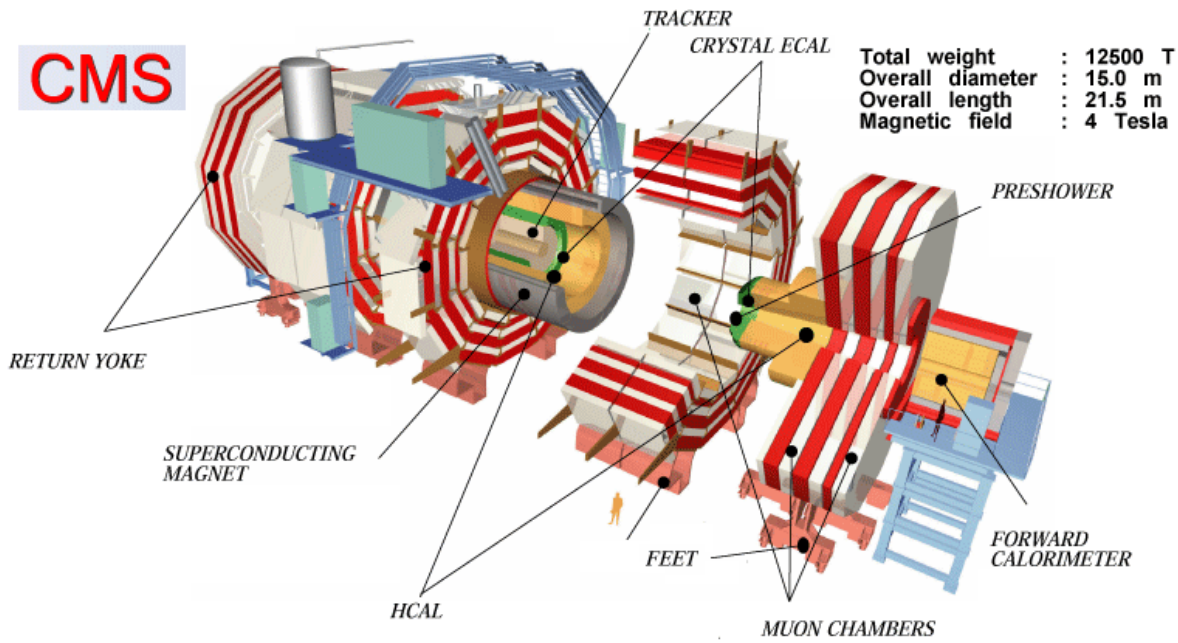

Figure 3-3. A diagram of CMS, showing its main detecting systems [31].

3.2.2 The Tracker

The CMS tracker is constructed to measure charged particles trajectories accurately. It was designed to give an excellent spatial resolution for all outgoing charged particles from the interaction region. Due to the nature of the beam, the tracker should combine a good spatial resolution with high performance in terms of vertex reconstruction. This needs to be done with minimum material budget to reduce effects such as multiple scattering and bremsstrahlung. The tracker is made of two subsystems: The pixel tracker and the silicon stripes. A brief description is given for both systems below:

3.2.2.1 Pixel tracker

The pixel detector is the closest tracker system to the interaction point. It covers a region of $|\eta| < 2.5$ and is organized in the form of 3 coaxial barrels that are $53\ \mathrm{cm}$ long with radii 4.4, 7.3 and $10.2\ \mathrm{cm}$. The caps are covered by two disks on each side at $\mathrm{z} = 34.5$ and $46.5\ \mathrm{cm}$. They are divided into pixels, each has an area of $100 * 150\ \mu\mathrm{m}^2$ giving around 44 million channels. At the end-caps, the pixels are rotated by 20° to

improve hit resolution. The pixel detector is able to provide single point hits with a spatial resolution of about $10\ \mu\mathrm{m}$ in the $\mathrm{r}-\phi$ space and $15-20\ \mu\mathrm{m}$ in the z direction.

3.2.2.2 Silicon strip tracker

The Silicon Strip Tracker is the outer part of the silicon tracker. It covers the pseudorapidity region $|\eta| < 2.5$, where radiation levels and track density are reasonable to operate silicon detectors. The strip tracker is divided to four subsections: The Tracker Outer Barrel (TOB), the Tracker Inner Barrel (TIB), the Tracker Inner Disk (TID), and the Tracker End Caps (TEC). Each subsection has layers of modules, where each module contains a set of silicon strip sensors. The dimensions of the strips vary between $117\ \mathrm{mm} * 64\ \mathrm{mm}$ for the inner modules and $190\ \mathrm{mm} * 96\ \mathrm{mm}$ for the outer ones.

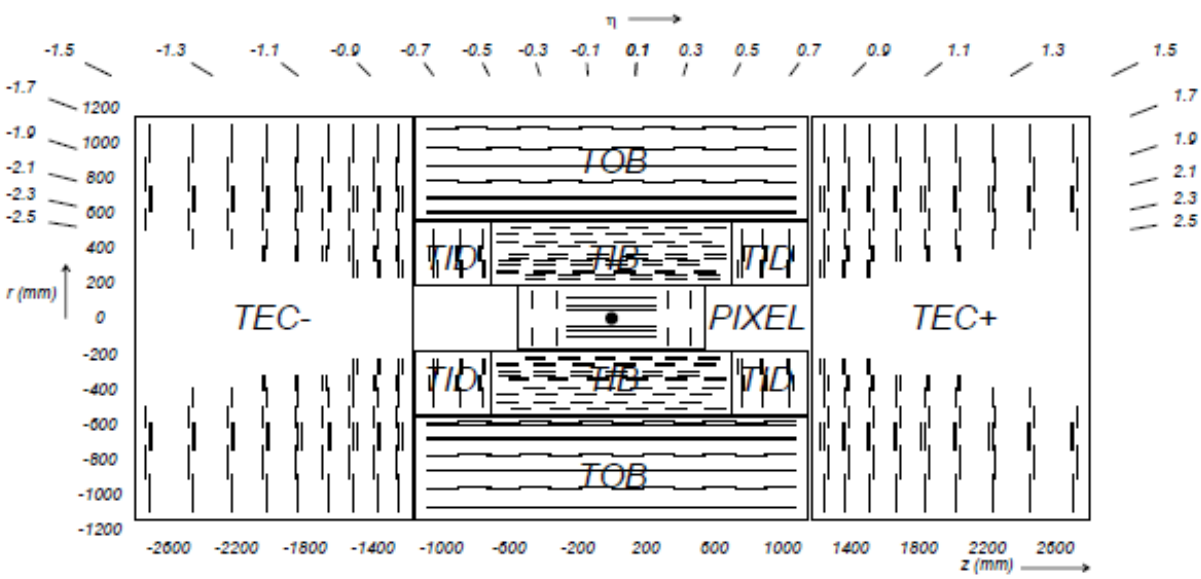

Figure 3-4. A longitudinal section view of the tracking system at CMS [32].

The resolution of the tracking system is around $1-2\%$ in $\mathrm{p_T}$ at $\mathrm{p_T} = 100\ \mathrm{GeV}$. The main factor that reduces the resolution is multiple scattering. At $\eta = 0$, there is a small decrease in efficiency due to the gaps in the pixel ladders.

3.3 Calorimeters

Calorimeters are an important detection system used for measuring the energy and position of a particle by its total absorption [33]. Calorimeters can be classified according to their structure: homogeneous or sampling. A homogeneous calorimeter can be used to detect photons by Cerenkov radiation emitted by e+e- pairs created in the Coulomb fields of the nuclei. Alternatively, a sampling calorimeter can be constructed from separate layers of an absorber and detector. CMS has both types of calorimeters: the homogeneous one is called The Electromagnetic Calorimeter (ECAL), and the sampling one is referred to as The Hadronic Calorimeter (HCAL).

3.3.1 The Electomagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) is one of the calorimetric systems of CMS. It is designed to measure photon and electron energies accurately. It is made from Lead-Tungestate scintillating crystals ($\mathrm{PbWO_4}$). There are 61200 crystals in the barrel reagion and 14648 at the end caps. The crystals at the barrel have a total radiation length of $\mathrm{X} = 25.8\mathrm{X_0}$ and $22\ \mathrm{mm} * 22\ \mathrm{mm}$ frontal cross section. The collection of the light is performed with silicon Avalanche Photo Diodes (APD). At the end caps, the crystals have $24.7\ \mathrm{mm} * 24.7\ \mathrm{mm}$ frontal cross section with a smaller radiation length $\mathrm{X} = 24.7\mathrm{X_0}$, in order to balance the existance of the pre-shower system in front of the end caps ($1.65 < |\eta| < 2.6$).

The energy resolution σ_E can be expressed as the square sum of three independent terms:

$$\left(\frac{\sigma_\mathrm{E}}{\mathrm{E}}\right)^2 = \left(\frac{\mathrm{S}}{\sqrt{\mathrm{E}}}\right)^2 + \left(\frac{\mathrm{N}}{\mathrm{E}}\right)^2 + \mathrm{C}^2 \ . \tag{3–5}$$

The first term in the right side of Equation (3–5) contains the stochastic term S and describes the effects of the fluctuations in photon statistics and shower containment.

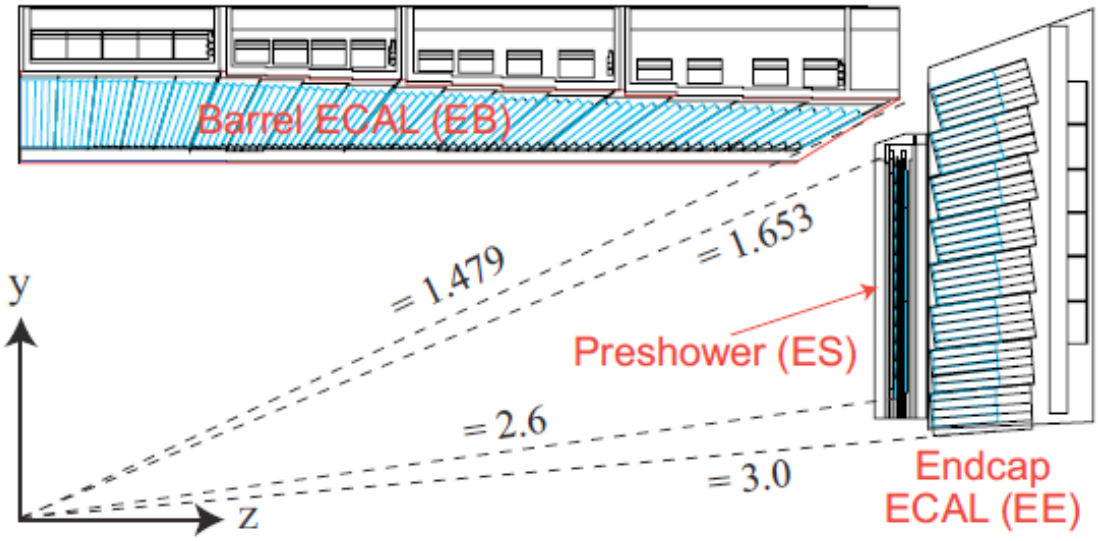

Figure 3-5. A longitudinal view of a quadrant of CMS, showing the ECAL system [34].

The N in the second term refers is proportional to photo-electric irregularities alongside the crystals, and C is a constant term affected by electronic noise and pile-up.

Equation (3–5) is valid up to $500\ \mathrm{GeV}$. After that, the shower leakage from the rear of the calorimeter becomes too significant and this parametrization of the resolution is no longer accurate.

3.3.2 The Hadronic Calorimeter

The Hadronic Calorimeter needs to fulfill two conditions in its design: it needs to possess enough density and thickness to contain the hadronic shower, and it should not be made with ferromagnetic materials since it is placed inside the superconducting magnet. HCAL is made from a sampling calorimeter with a $3.7\ \mathrm{mm}$ plastic scintillator alternated with a $5\ \mathrm{cm}$ thick brass plate absorber. Similar to ECAL, the HCAL barrel covers the region of $|\eta| < 1.48$ and the end cap covers $1.48 < |\eta| < 3.0$. The barrel thickness is around $79\ \mathrm{cm}$, with $6.5\ \lambda_0$ of brass. Since this is considered too thin, another layer equal to $3.5\ \lambda_0$ is added beyond the solenoid. To cover more forward regions ($3.0 < |\eta| < 5.0$), which improves $\mathrm{E_T^{miss}}$ measurements, a forward calorimeter is located outside the magnet yoke, at $\pm 11\ \mathrm{m}$ from the interaction point. It is made of layers of

quartz and steel and known as the HF, see Figure 3-6. The energy resolution of HCAL is:

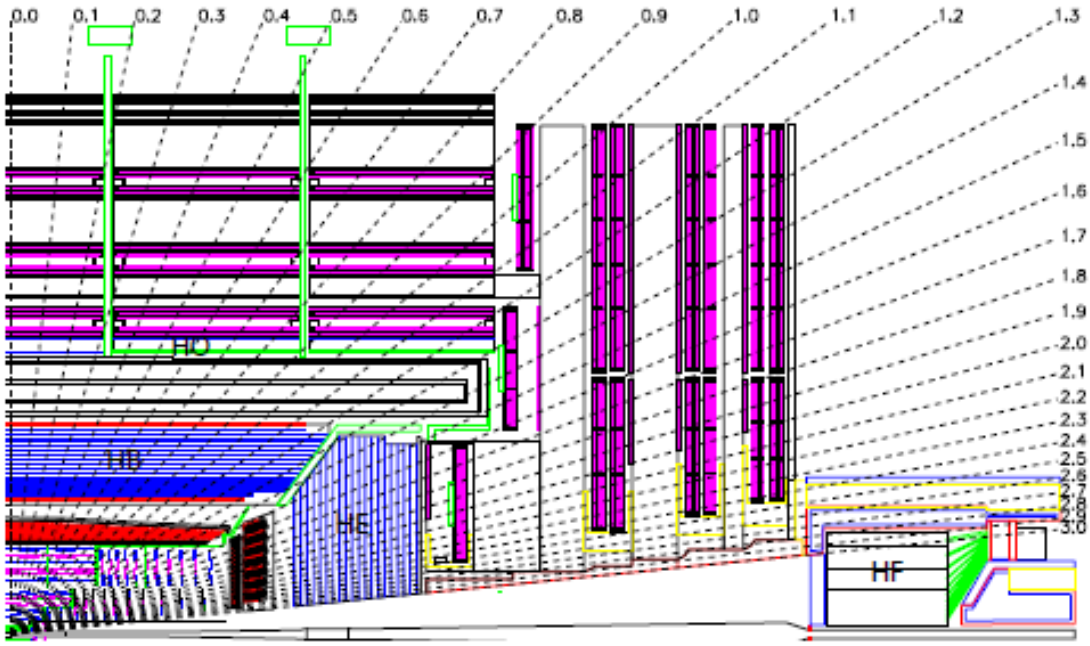

Figure 3-6. A longitudinal view of a quadrant of CMS, showing the HCAL system [32].

$$\left(\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\right)^2 = \left(\frac{140.2\%}{\sqrt{\mathrm{E}}}\right)^2 + (4.7\%)^2 \quad . \tag{3–6}$$

The first term represents the stochastic fluctuation in the signal. The second term is for fluctuations of the shower accompanying the signal. In the HCAL, the resolution gets sensibly worse for $|\eta|$ larger than 1.4. This is mainly due the inactive material in the detector (i.e. service cables). The performance of the forward calorimeter (HF) can be divided into hadronic and electormagnetic components:

$$\left(\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\right)^2_{\mathrm{hadronic}} = \left(\frac{182\%}{\sqrt{\mathrm{E}}}\right)^2 + (9\%)^2 \quad , \tag{3–7}$$

$$\left(\frac{\sigma_{\mathrm{E}}}{\mathrm{E}}\right)^2_{\mathrm{electromagnetic}} = \left(\frac{138\%}{\sqrt{\mathrm{E}}}\right)^2 + (5\%)^2 \ . \tag{3–8}$$

3.4 Forward Detectors

3.4.1 Centauro and Strange Object Research

The Centauro and Strange Object Research (CASTOR) is a quartz-tungsten sampling calorimeter installed at $14.4\ \mathrm{m}$ from the interaction point. It covers $5.2 < |\eta| < 6.6$ and it serves many physics analyses dealing with forward hadrons in proton-proton and ion-ion interactions.

3.4.2 Zero Degree Calorimeter

The Zero Degree Calorimeter (ZDC) is a combanation of sampling quartz-tungsten electromagnetic and hadronic calorimeter. Two identical ZDCs are placed $140\ \mathrm{m}$ away from the interaction point from each side of the detector. It will detect photons and neutrons with $|\eta| < 8.3$. ZDC is used in many studies, such as centrality studies in heavy-ion collisions.

3.5 The Muon System

While hadronic background is mainly contained within the calorimeter, muons are able to propagate with minimal interaction with the detector. Muons therefore provide a strong indication of a signal event over minimum bias background and are therefore prime candidates for triggering signal.

CMS employs a high performance muon system for fast identification of muons that also gives good momentum resolution. The muon system at CMS consists of 3 subsystems: Drift tubes (DT) detectors on the barrel, Cathode Strip Chambers (CSC) within the yoke plates, both of these subsystems cover a region of $|\eta| < 2.4$. And in the region of $|\eta| < 2.1$, redundancy is provided by Resistance Plate Chambers (RPC), which have an excellent time resolution and a fast response that balances its relatively limited

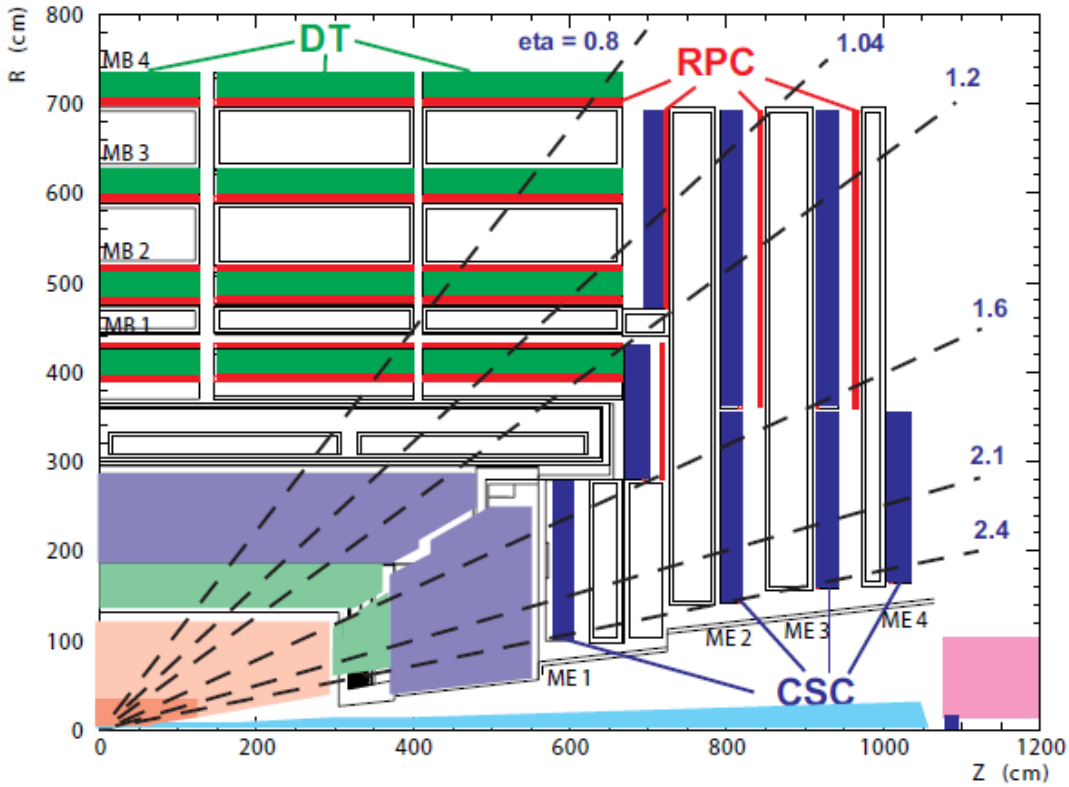

Figure 3-7. The layout of one quarter of the CMS muon system [34].

spatial resolution. They are used mainly to complement DT and CSC measurements and to identify bunch crossings.

Using the muon system in conjunction with the tracking system helps in improving the momentum resolution by an order of magnitude for both regions of low $\mathrm{p_T}$ (where tracks are affected by multiple Coulomb scattering), and high $\mathrm{p_T}$ (where the sagitta of the muon track can be measured before and after the solenoid).

3.6 Triggers

A trigger system aims to accept all useful events and to reject most backgrounds. This is crucial at an experiment such as the LHC, where the beam frequency is $400\mathrm{M\ Hz}$. Such high frequency leaves only $25\ \mathrm{ns}$ for data readout and processing. It is technically challenging to read out, make a decision about the event, and store the large amount of data at this rate; not to mention selecting the signals that are relevant to

the event. The solution is to pipe the data to a strong processing system that takes few microseconds to make the decision to accept the event or not. Such selection process requires a good tiggering strategy that maximizes events with discovery potential. The trigger system at CMS has two stages: Level one triggers (L1) and High level triggers (HLT):

3.6.1 Level 1 Triggers

A Level 1 trigger is designed to perform very fast accept/reject decisions. L1 trigger decision is based on calorimeter towers and muon chamber information only. This reduces the rate of data to $10\mathrm{k\ Hz}$. After that, there are 3 layers of increasing complexity of L1 triggers: local (such as calorimeter towers), regional (combinations of local triggers), and global (which performs calculations for variables such as jets and total transverse energy $\mathrm{E_{Tsum}}$. See diagram 3-8 for more details of L1 triggering stratefy at CMS.

3.6.2 High Level Triggers

The High Level Trigger is a software system implemented in a cluster of commercial processors at CMS. It performs the readout of the front-end electronics for events accepted at L1-trigger. There is no limitation on the number of virtual trigger levels or to the algorithms employed except the CPU time. At CMS, the HLT trigger strategy is the traditional multi-level trigger systems, where the selection process is optimized by rejecting uninteresting events as quickly as possible. With this in mind, each trigger path consists of a sequence of software modules with increasing complexity and physics sophistication. Each module fulfills a well defined task such as reconstruction, intermediate trigger decisions or the final trigger decision for that path. After the HLT the event rate will be in the order of $100\ \mathrm{Hz}$, which is manageable within the current storage resources at CMS.

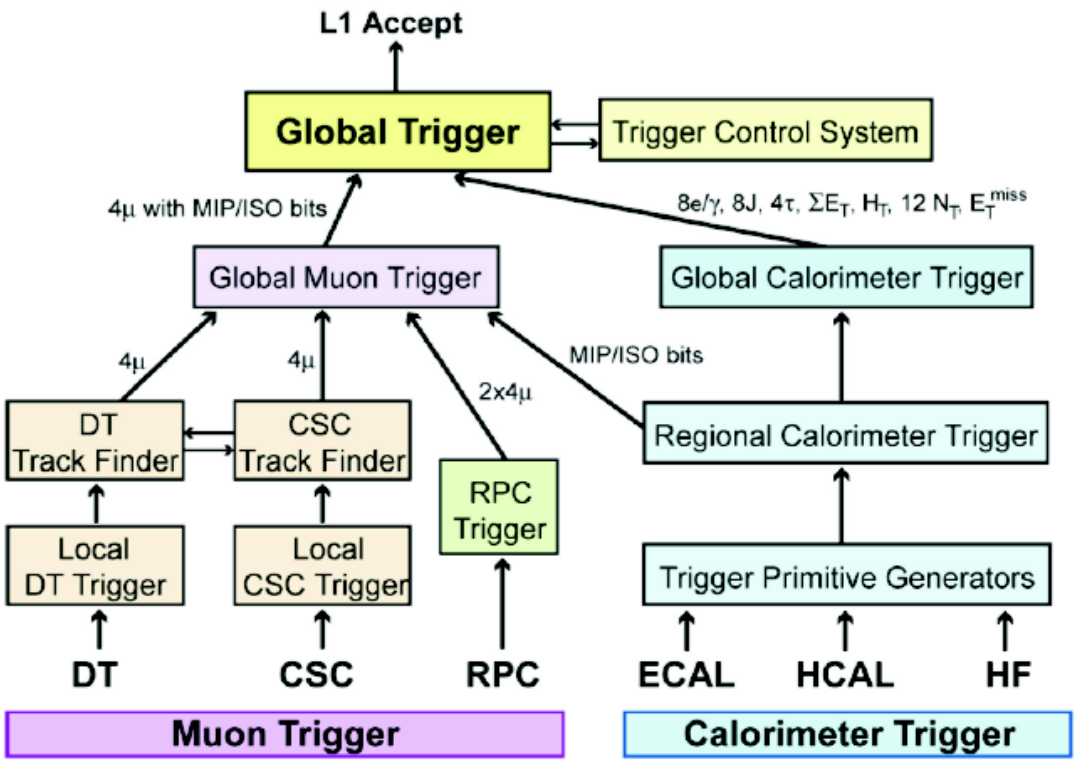

Figure 3-8. An overview of CMS Level-1 triggers [34].

CHAPTER 4
TRACKING AT CMS

CMS sees an event as a large amount of electronic signals in its systems. The next step is processing this information to reconstruct the particles history after the collision. Since this analysis depends mostly on tracks, it is befitting to introduce the basics of track reconstruction in CMS in some detail.

4.1 Track Reconstruction

Track reconstruction is traditionally divided into two separate subtasks: track finding and track fitting [35]. Track finding is the process of determining the subset of measurements in the tracking system that belongs to the same track. Track fitting extrapolates the measurements from the track finding stage to estimate a set parameters that identifies the track. Additionally, the quality of the track candidates is evaluated at this phase and a decision is made if it should be accepted as a real track or not.

This division between track fitting and track finding lasted until the ’80s, when Kalman filter [36] was invented. It can be viewed as a statistically optimal refinement of track following [37]. Generally speaking, a track at any given surface of the detector can be described by 5 track parameters: 2 for position, 2 for direction, and one for curvature or momentum, also known as the state vector of the system. A Kalman filter consists of a series of alternating predicton and filter steps. In the prediction step, the state vector is extrapolated to the next detector surface. In the filter step, the extrapolated state vector is updated by taking a weighted mean of the measurement.

Shortly after the invention of the Kalman filter, it was realized that this procedure can be used for both track finding and track fitting. Due to the uniform magnetic field

at CMS, the outgoing charged particles from the point of interaction will be moving in a helical path. The 5 parameters used by CMS to describe such path are [1] :

- $\mathrm{p_T}$, the transverse momentum.
- $\cot(\theta) = \mathrm{p_z}/\mathrm{p_T}$, the dip angle that complements the angle between $\vec{\mathrm{p}}$ and $\vec{\mathrm{p_T}}$.
- ϕ, the azimuthal angle of the momentum vector at the impact point.
- $\mathrm{d_0}$ and $\mathrm{d_z}$, the transverse and longitudinal distance with respect to the nominal vertex. They are set to have the closest distance to it.

The Quality of the tracks is given in terms of χ^2 of the track fit, the number of degrees of freedom of the fit, the number of hits n used in the fit, and by the number of gaps in the measurement sequence $\mathrm{n_{losthits}}$.

The following summarizes the general steps for track reconstruction in CMS [35]:

- Seed generation, where the seed consists of few measurements, mostly in the pixel part of the tracker.
- Local track finding beginning from a seed. At the LHC this is done through algorithms based on generalized versions of Kalman filter.
- Track fitting, which involves procedures such as removing track candidates that have too many common measurements (trajectory cleaning). For example Gaussian-sum filter (GSF) is used for electron track fitting.
- Post-processing; refitting, ambiguity resolution, etc.

It is worth mentioning that the muon system has a dedicated strategy to handle its tracks. This is because it has a reduced magentic field that is generally not as well behaved as the field in the central region, it is important to account for the passage of muons through more material. CMS uses combinatorial Kalman filter for muon tracking.

[1] This is not the only choice to describe a track, For example one can use :$\mathrm{q/p}$, $\mathrm{dx/dz}$, $\mathrm{dy/dz}$, $(\mathrm{x,y})$ and $\mathrm{sign(z)}$ to uniquely identify a track

4.2 Track Quality

After the process of fitting the tracks and parameterizing them, the tracking software needs to determine the level it trusts that these tracks correspond to physical tracks. The result of such test is given in terms of track quality. The tracks that have low value of quality are considered fake or ghost tracks. The criteria applied in rejecting ghost tracks are [38]:

- χ^2/n.
- $\mathrm{d_0}$, the track distance from the beam spot.
- The track Δ_z to the position closest to HLT primary vertex.
- The $\mathrm{d_0}/\delta\mathrm{d_0}$ where $\delta\mathrm{d_0}$ is the measured error in the transverse beam spot postion.
- The $\Delta_\mathrm{z}/\delta\mathrm{d_z}$ longitudinal compatibility with the closest HLT vertex.

In order to adapt the track quality cuts to the expected track resolution for the vertex association, the resolutions on the track measured $\mathrm{d_0}$ and $\mathrm{d_z}$ have been parametrized as:

$$\sigma_{(\mathrm{d_0},\mathrm{d_z})}(\mathrm{p_T}) = \mathrm{a} + \frac{\mathrm{b}}{\mathrm{p_T}[\mathrm{GeV/c}]} \ , \tag{4–1}$$

where a and b are configurable parameters that can be set to the nominal values $\mathrm{a} = 30\ \mu\mathrm{m}$ and $\mathrm{b} = 100\ \mu\mathrm{m}$ or tightened to reduce the fake rate at low $\mathrm{p_T}$: The dependencies of the optimal cuts with the track $\mathrm{n_{layers}}$ (number of layers crossed), $\mathrm{p_T}$ and η have been approximated with the following formulas:

- $\chi^2/\mathrm{n} < \alpha_0 \mathrm{n_{layers}}$.
- $|\mathrm{d_0}| < (\alpha_1 \mathrm{n_{layers}})^{\mathrm{x_1}} \sigma_{\mathrm{d_0}}(\mathrm{p_T})$.
- $|\Delta_\mathrm{z}| < (\alpha_2 \mathrm{n_{layers}})^{\mathrm{x_2}} \sigma_{\mathrm{d_z}}(\mathrm{p_T})$.
- $|\mathrm{d_0}|/\delta\mathrm{d_0} < (\alpha_3 \mathrm{n_{layers}})^{\mathrm{x_3}}$.
- $|\Delta_\mathrm{z}|/\delta\mathrm{d_z} < (\alpha_4 \mathrm{n_{layers}})^{\mathrm{x_4}}$.

Where i, $\mathrm{x_j}$ (with $\mathrm{i,j} \in \{1,2,3,4\}$ are configurable parameters). CMS [38] has adopted few track selection scenarios to supress fake tracks while keeping high efficiency. CMS collaboration has implemented three different thresholds: loose, tight and high Purity. The first collection of tracks are all the tracks reconstructed from the algorithm of tracking. Tight selection and $\mathrm{highpurity}$ selection only differ in impact parameter cuts, the latter is stricter due to different parametrizations in Equation (4–1), for high Purity the adopted option was $\mathrm{a} = 30\ \mu\mathrm{m}$ and $\mathrm{b} = 10\ \mu\mathrm{m}$. In the following, we study how charged particles produced with pseudorapidities within the tracker coverage ($|\eta| < 2.5$) and transverse momenta $\mathrm{p_T}$ above $0.5\ \mathrm{GeV/c}$ can be reconstructed and qualified as a $\mathrm{highpurity}$ track.

In order to increase the performance of CMS tracking, Iterative tracking is applied in selecting the tracks. After removing hits used in the previous iteration, where a new seed and tracking parameters are used in each iteration. In the 5 iterations, there are changes in the seeding layer (pixel triples, doublet, or other combinations including the strip tracker), $\mathrm{p_T}$, $\mathrm{d_0}$, and $\mathrm{d_z}$.

4.3 Primary Vertex Reconstruction

Once tracks are reconstructed, dedicated algorithms are applied to estimate the primary vertex position. This process, known as vertex reconstruction, typically involves two steps: Vertex finding, where clusters of tracks originating from the same vertex are grouped together as vertex candidates, and vertex fitting, where the position of the primary vertex is computed.

4.3.1 Vertex Finding

To speed-up the primary vertex finding process, pixel tracks from pixel hit triplets can be used to efficiently find the vertex positions. In off-line analyses, where timing is not an issue, the full information of the reconstructed tracks and their corresponding covariance matrices are used. This off-line method [39] starts with a preselection cut based on the impact parameter significance $\mathrm{d0}/\sigma_{\mathrm{d0}}$. This cut rejects secondary vertices

and fake tracks, and reduces the computation time. In a second step, the selected tracks are extrapolated to the beam line, and grouped according to their separation in the z direction in order to form primary vertex candidates. A limit is set on the maximum separation in z between the two successive group of tracks to the same primary vertex candidate. Eventually, the reconstructed primary vertices are sorted in decreasing order of hardness, defined by the scalar sum of the tansverse momentum of each track squared.

$$\Sigma_{\mathrm{i}=1}^{\mathrm{N}}(\mathrm{p}_{\mathrm{T_i}})^2 \quad . \tag{4–2}$$

4.3.2 Vertex Fitting

The set of tracks associated with the primary vertex fitting algorithms are used to compute the best estimate of the vertex parameters such as its position and covariance matrix, as well as indicators of the fit quality [40]. Vertex fitting in CMS is performed using a statistical method known as the Adaptive Vertex Fitter [40]. It is considered superior to least-square Kalman filters due to its better handling with mis-associated tracks (outliers). In the adaptive vertex fit, each track attached to a vertex is assigned a track weight between 0 and 1 based on its compatibility with the common vertex. The tracks with larger distance to the vertex position are down-weighted significantly, which makes the algorithm robust against outliers. The number of degrees of freedom is defined as $2\Sigma_{\mathrm{n}}^{\mathrm{Tracks}}\mathrm{w_i} - 3$, where $\mathrm{w_i}$ is the weight of ith track. It is thus strongly correlated to the number of tracks compatible with the primary interaction region. Consequently, the number of degrees of freedom of the vertex can be used to select real proton-proton interactions.

4.4 Track Selection for the UE Analysis and Efficiencies

The track selection criteria follows the strategy employed by previous UE analyses at CMS [27]. To measure the UE, we need to find the tracks associated with p-p

collisions (i.e. primary tracks). This strategy tries to minimize all tracks coming from secondary interaction and fake tracks. In addition to this criteria, there should be emphasis on good track resolution. This is because most of all the observables will be presented as a function of $\mathrm{p_T}$ of the leading track. A track is considered primary if it has $\mathrm{c}\tau < 1$. Reconstruction efficiency is estimated by comparing them to associated hits that match the hits expected from a charged particle passage vs the clusters reconstructed in a detector. A track is considered associated if the shared hits are more than 75%.

In order to quantify the performance of a track finding algorithm, a MC simulation of a track sample, where all the tracks are labeled, is used. We use the efficiency ϵ_{P} to measure it. It is defined as the number of associated tracks the algorithm can find divided by the total number of simulated primary tracks. To be of real use, efficiency needs to be accompanied with the purity $\mathrm{P_p}$ of the sample of the found tracks, which is the number of associated tracks with the correct label divided by the total number of selected tracks. Efficiency is more important than purity, since a near perfect track with few wrong hits can very likely be corrected, while a track not found during the track search is lost forever [37]. Studying efficiency in real data is much more difficult, and it usually involves scanning an event sample visually or using independent sources such as scintillators. See Figure 4-1 for these values for the used tracking algorithm, the different MC labels represent different track cuts, a topic of the next section.

4.4.1 Track Cuts

To select the primary track in our sample, a set of cuts have been introduced:

- $|\eta| < 0.8$,
- $\mathrm{p_T} > 0.5\ \mathrm{GeV/c}$,
- $\frac{\mathrm{d_0(vtx)}}{\sigma_{\mathrm{d_0}}} < 3$,
- $\frac{\mathrm{d_z(vtx)}}{\sigma_{\mathrm{d_z}}} < 3$.

The last two cuts are used to select the tracks associated to a primary vertex. On top of the base selection, some quality cuts are required to reduce the fake contribution and to reject tracks with poor momentum measurement. Three selection choices are developed:

- UE1: base selection with High Purity and $\frac{\sigma(\mathrm{p_T})}{\mathrm{p_T}} < 10\%$,
- UE2: base selection with $\mathrm{N_{layers}} \geq 4$ (the minimum number of strip layers crossed by tracks), $\mathrm{N_{pixellayers}} \geq 2$ (minimum number of pixel layers) and $\frac{\sigma(\mathrm{p_T})}{\mathrm{p_T}} < 5\%$,
- UE5: base selection with High Purity and $\frac{\sigma(\mathrm{p_T})}{\mathrm{p_T}} < 5\%$,

UE5 is chosen as reference while UE1 and UE2 are used as a cross-check. The only difference between UE1 and UE5 is the looser cut on $\sigma(\mathrm{p_T})/\mathrm{p_T}$, while UE2 does not use the $\mathrm{highpurity}$ quality, but instead relies on the minimum number of crossed layers. The performance of the UE5 selection

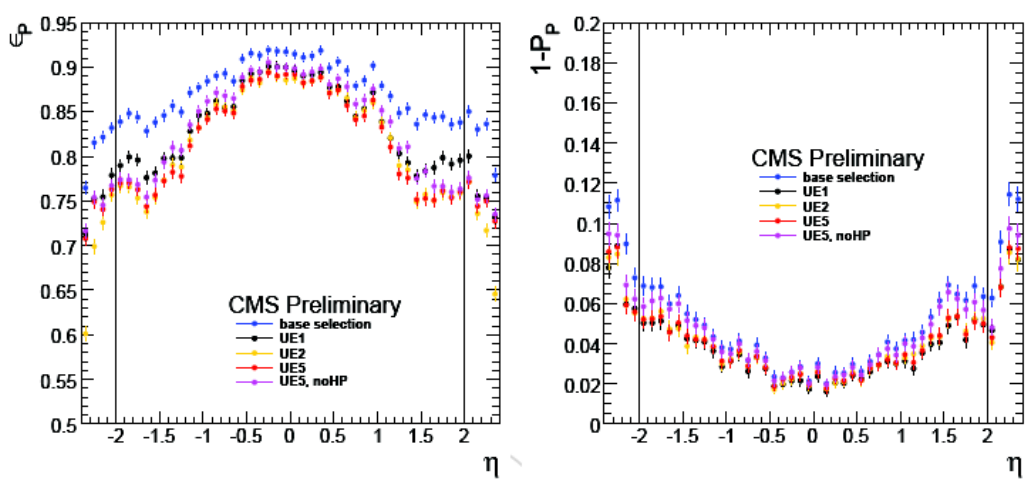

Figure 4-1. The efficiency and background contribution for the three track selection criteria [41].

The left panel of Figure 4-1 shows the efficiency, ϵ_P, to be close to 90% in the central region, dropping to 75% for $|\eta| < 2$, and in the right panel of Figure 4-1 we observe the background contribution. As expected the degradation in efficiency and the increase in the fake rate are correlated, in the transition barrel-endcap regions of the

tracker, which is characterized by a larger material budget. The result is an increase of the probability of inelastic nuclear interactions of the primary particles, which converts them to two or more secondary particles. The 3% of non-primary tracks passing the selection are attributed to combinatorial background (2%) and secondary tracks (1%) resulting from $\mathrm{K_S}$ and Λ decay [27].

4.4.2 Transverse Momentum Uncertainty

As it was explained earlier, there is a need to seek tracks with as much precision as possible, bearing in mind not to tighten the selection criteria to the limit of eliminating many good tracks. Figure 4-2 shows the variable $\sigma(\mathrm{p_T})/\mathrm{p_T}$ for data and different Monte Carlo models. In the plot, the second bump is due to primary particles creating less hits (i.e. electron-hole pairs) than expected because of nuclear interactions or decays. In particular, these tracks have hits in the pixel detector only, giving a wrong estimations of the transverse momentum.

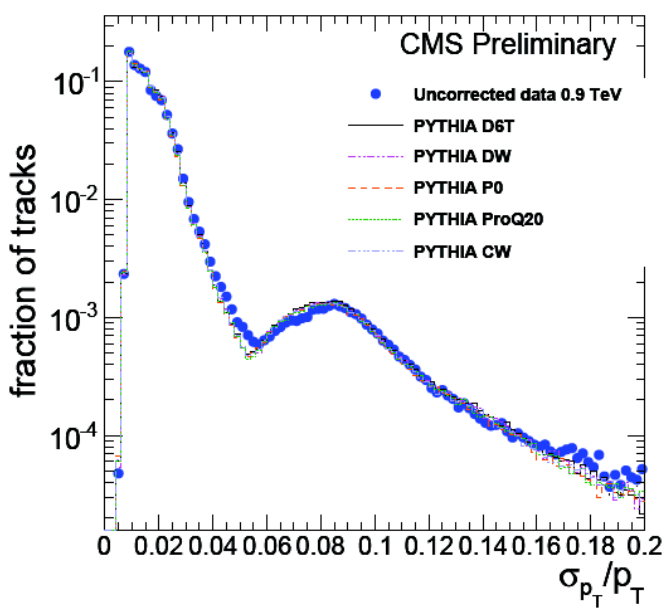

Figure 4-2. The fraction of tracks vs. $\sigma(\mathrm{p_T})/\mathrm{p_T}$ at $0.9\ \mathrm{TeV}$ [41].

similar results were observed for $7\ \mathrm{TeV}$ [42].

CHAPTER 5
THE UE ANALYSIS STRATEGY

5.1 Overview

In Chapter 5 we will list the data samples used in the analysis and briefly describe the methods applied in extracting the events deemed useful for the UE studies from the background. These methods are either applied during data acquisition such as triggering, or after the data have been stored, such as event selection. Both will be discussed as well as the MC samples that were used in detector studies.

5.2 Triggering Strategy

The data streaming out of the CMS detector needs to be split into data sets to take advantage of the parallel CMS computing model for data processing. Each set is called a Primary Data set (PD) [43]. The used PDs are based on Zero Bias and Minimum bias trigger configurations:

- The Zero Bias PD is based on the Beam Pick-up Timing eXperiment (BPTX) [44]. This system is designed to provide precise information regarding the timing of the LHC beam. It has a time resolution of the order of $2 * 10^{-10}$ s [44].
- The Minimum Bias PD is based on the Beam Scintillation Counter (BSC). It measures the relative rates of background particles and collision products, both entering and exiting CMS. It is stationed 10.86 meters away from the interaction point from both ends with time resolution of the order of 3 ns [45].

In Table 5-1 the technical trigger bits dealing with BPTX and BSC are reported. A trigger strategy with high efficiency with collision events was needed, it depends on the following trigger bits:

- Bit 40: it requires a coincidence inside a 20 ns window between two BSC stations with at least one hit in each station.
- Bit 41: it requires a coincidence within 20 ns between the two BSC stations with at least two hits in each station.
- Bits 36, 37, 38, 39 (the beam halo triggers): they require hits in both BSC stations to be separated in time by 73 ± 20 ns, corresponding to the time of flight between the two BSC stations.

- It is also required to have the beam present (bit 0), indicating a filled bunch from both beams crossing the interaction point at the same time based on the corresponding BPTX signal.

The trigger requested is a combination of the bits described above:

$$0 \cap 40 \cap !(36 \cup 37 \cup 38 \cup 39) \ . \tag{5–1}$$

Table 5-1. Table listing the trigger bins used for the BSC and BPTX triggers.

Bin	Trigger	Trigger Name	Trigger Details
0	BPTX	L1TechBPTXplusANDminus	$BPTX + andBPTX-$
32	BSC	L1TechBSCminBiasinnerthreshold1	$BSCminbias : hits + zside >= 1andhits - zside >= 1$
33	BSC	L1TechBSCminBiasinnerthreshold2	$BSCminbias : hits + zside >= 2andhits - zside >= 2$
34	BSC	L1TechBSCminBiasOR	$BSCOR : thereisatleastonehit, anywhereintheBSC$
35	BSC	L1TechBSCHighMultiplicity	$BSChighmultiplicity : allcoincides$
36	BSC	L1TechBSChalobeam2inner	$BSChalo : beam2inner$
37	BSC	L1TechBSChalobeam2outer	$BSChalo : beam2outer$
38	BSC	L1TechBSChalobeam1inner	$BSChalo : beam1inner$
39	BSC	L1TechBSChalobeam1outer	$BSChalo : beam1outer$
40	BSC	L1TechBSCminBiasthreshold1	$BSCminbias : hits + zside >= 1andhits - zside >= 1$
41	BSC	L1TechBSCminBiasthreshold2	$BSCminbias : hits + zside >= 2andhits - zside >= 2$
42	BSC	L1TechBSCsplashbeam1	$BSCsplashtrigger : beam1, threshold = 2; innerringhitson - zside >= 2$
43	BSC	L1TechBSCsplashbeam2	$BSCsplashtrigger : beam2, threshold = 2; innerringhitson + zside >= 2$

In this analysis it is important to determine if the trigger efficiency is compatible between data and different MC models. The efficiency of the trigger has been computed with event candidates taken from the Zero Bias sample. The event selection corresponding to this analysis is required to have at least one high purity track with $\mathrm{p_T}$ above 0.5 GeV /c in the tracker acceptance region $|\eta| < 2$. The trigger efficiency is defined as the ratio of the number of events that pass the event selection and fire the trigger over the number of events that pass the event selection.

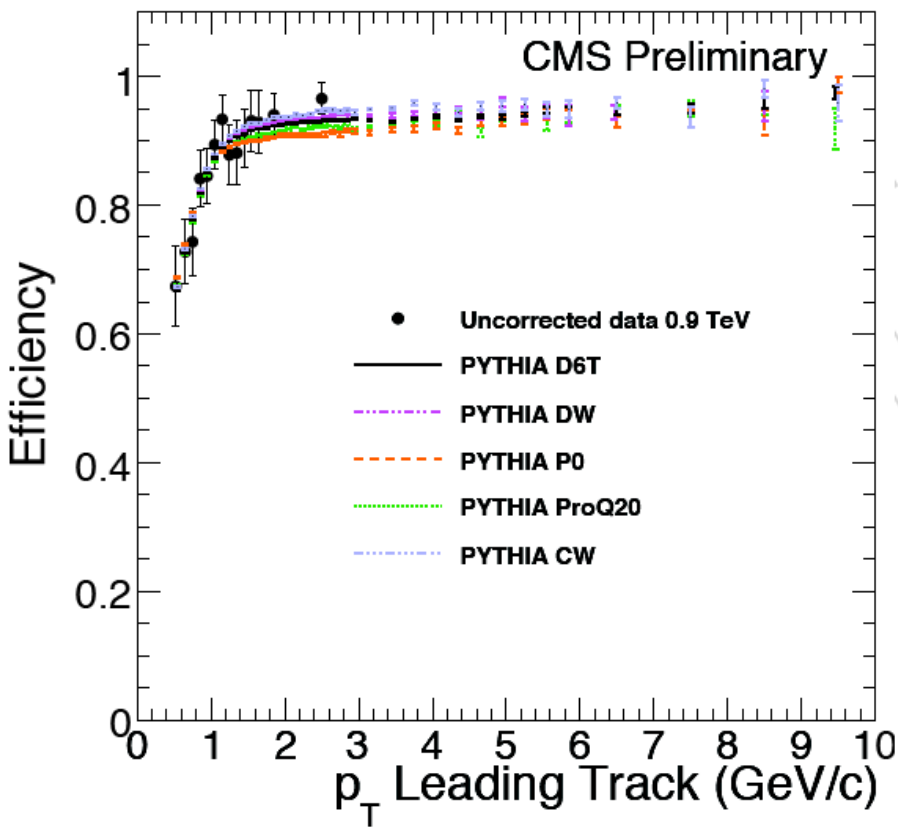

Figure 5-1. BSC trigger efficiency vs the leading track momentum at $0.9\ \mathrm{TeV}$ [41].

Figure 5-1 shows that the MC efficiency of the BSC trigger selection agrees well with the efficiency derived from the Zero Bias data. The efficiency is close to $\%95$ for a

track $\mathrm{p_T} > 2\mathrm{GeV/c}$. There is a good agreement between the data and the reference MC simulation.

Similar trigger choices are implemented at $7\ \mathrm{TeV}$, with the same PD convention: Zero Bias PD that depends on the BPTX trigger, and Minimum Bias PD that uses the BSC trigger. For $7\ \mathrm{TeV}$, the following Minimum Bias bits have been used:

- Bit 40.
- Bits 36, 37, 38, 39.
- Bits 42 (43), BSC splash trigger for beam 1 (2). The threshold is set at 2, asking for inner ring hits.

The logic implemented considering the numbering scheme over-reported is the following:

$$0 \cap 40 \cap !(36 \cup 37 \cup 38 \cup 39) \cap !((42 \cap !43) \cup !(43 \cap !42)) \ . \tag{5–2}$$

Bits 42 and 43 were not available in the first collisions at 900 GeV, they allow us to improve the selection of true collision between the beams.

Our analysis is based on the following runs:

- Pre-1E29 (Runs 132440-135807): 1 PD
- 1E29 (Runs 135808-140041): 8 PDs
- 1E30 (Runs 140042-141949): 9 PDs

The triggering strategy was modified due to code migration (from $\mathrm{CMSSW\_3\_5\_X}$ to $\mathrm{CMSSW\_3\_6\_X}$), there was a shift to using HLT trigger: $\mathrm{HLT\_MinBiasPixel\_SingleTrack}$ is applied, where we require the existence of at lest one track in the pixel detector with $\mathrm{p_T} >= 0.2\mathrm{GeV}$. This trigger is seeded by $\mathrm{L1\_BscMinBiasOr\_BptxPlusOrMinus}$ (That is firing either bit 0 or bit 40). This path maintains the high efficiency of the first choice.

5.3 Data and MC Samples

This analysis uses data samples covering two center of mass energies:

- $0.9\ \mathrm{TeV}$ data sample collected in 2009. The data sample used has 11.0 M of triggered events, for an integrated luminosity of about $1\mu\mathrm{b}^{-1}$.
  The following data sample was used:

  – MinimumBias/Commissioning10-Jun14thReReco-v1/RECO

- $7\ \mathrm{TeV}$ data sample, The data used in this analysis have been taken from March to May of 2010. The analysed data correspond to an amount of $1\mathrm{nb}^{-1}$ of integrated luminosity. Due to the limited pre-scaling of the Minimum Bias trigger this correspond to a total statistics of 28.5M events triggered. The data at $7\ \mathrm{TeV}$ have 4 different samples (named from 1 to 4). They correspond to slightly different running conditions at the LHC.
  The following data samples are used:

  – Data set I : /MinimumBias/Commissioning10_Jun14thReReco-v1/RECO

  – Data set II : /MinimumBias/Run2010A_Jun14thReReco-v2/RECO

  – Data set III : /MinimumBias/Run2010A_Jul16thReReco-v1/RECO

  – Data set IV : /MinimumBias/Run2010A_PromptReco-v4/RECO

  As for the MC samples, there are two required distinctions: MC samples that represents the hadronic activity at the generator (GEN) level only, i.e. the output of simulating the physical processes of proton-proton collisions. And the MC samples that represent the physical process after passing a detector simulation (SIM). While we can easily generate GEN samples privately, the production of SIM samples is more restricted and we only used officially approved CMS samples:

  – MinBias-TuneZ1-900GeV-pythia6/Summer10-START36-V10A-v1

  – MinBias-TuneD6T-900GeV-pythia6/Summer10-START36-V10A-v1

  – MinBias-TuneD6T-7TeV-pythia6/Summer10-START36-V10-SP10-v1

  – MinBias-TuneZ1-7TeV-pythia6/Summer10-START36-V10-TP-v1

  These samples are used for experimental procedure validation, and to correct the data to the generator level. The amount of MC samples available varies between different tunes, but they are all above 10M events level, thus providing enough statistics.

5.4 Event and Track Selection

After passing through the trigger phase, an event needs to survive a set of criteria designed to reject background events. A beam scraping filter is applied to the triggered events. Then we require only one primary vertex to be among the selected events. This vertex needs to be valid (i.e. it should have different coordinates than the beam spot).

The case of vertex splitting was inspected and it was found not to give any significant bias in the analysis results [42]. The vertex should have at least 5 degrees of freedom and should be within $10\mathrm{cm}$ from the z direction of the beam spot, as shown in Figure 5-2 this cut is within 3σ of the gaussian fit average. Tables 5-2 and 5-3 show the percentage of events that survive the application of each one of the selection criteria. They also show the percentage of the surviving events for one of the MC tuens used in this analysis.

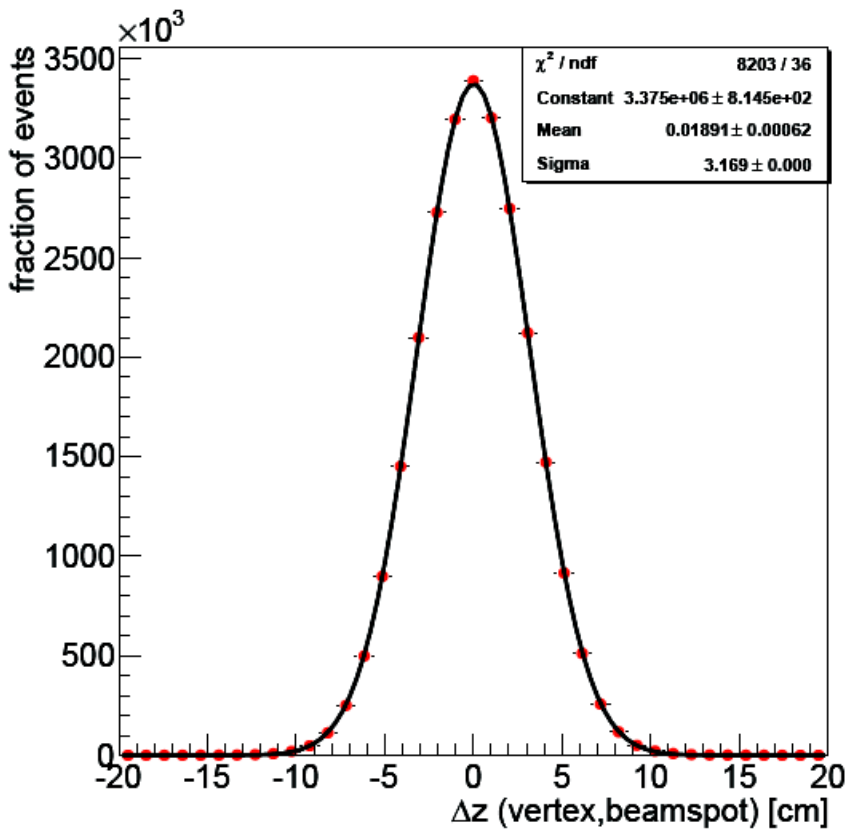

Figure 5-2. The difference in the z axis between the vertex and the beam spot.

Further quality controls with tighter cuts were needed to reduce the background coming from various sources, for each selected event the reconstructed track collection needs to be cleaned up from undesired tracks, namely secondary tracks (tracks resulting from the decay of unstable particles) and combinatorial background (fake

Table 5-2. The number of events surviving different event selection criteria at $0.9\ \mathrm{TeV}$.

Event Selection	Data	MC(Z1)
Triggered	$11.0\mathrm{x}10^6$	100%
+1 valid Primary Vertex	91.6%	90.5%
+ (+/- 10 cm) vertex window	86.7%	89.7%
+ vertex ndof larger than 4	85.6%	83.2%

Table 5-3. The number of events surviving different event selection criteria at $7\ \mathrm{TeV}$.

Event Selection	DataI	DataII	DataIII	DataIV	MC(Z1)
Triggered	$18.7\mathrm{x}10^6$	$33.2\mathrm{x}10^3$	$2.7\mathrm{x}10^6$	$11.4\mathrm{x}10^6$	$9.5\mathrm{x}10^6$
+1 valid Primary Vertex	95.3%	95.8%	95.8%	95.9%	93.8%
+ (+/- 10 cm) vertex window	99.6%	99.3%	89.3%	88.3%	99.6%
+ vertex ndof larger than 4	90.7%	90.6%	90.7%	90.7%	87.6%

tracks coming from erroneous reconstruction of a track from segments that belong to other tracks). The cuts applied are:

- $\mathrm{p_T} > 0.5\mathrm{GeV/c}$.
- $|\eta| < 0.8$.
- Track Quality: High Purity selection and $\sigma(\mathrm{p_T})/\mathrm{p_T} < 0.05$.
- Secondaries removal, non-primary particles resulting from decays of secondary interaction have large impact parameter with respect to the beam axis and are not associated with the position of the event vertex.
  - the longitudinal impact parameter significance $\mathrm{d(vtx)_0}/\sigma_{\mathrm{d0}} < 3$.
  - the transverse impact parameter significance $\mathrm{d(vtx)_z}/\sigma_{\mathrm{dz}} < 3$.
    The distributions of $\mathrm{d(vtx)_0}/\sigma_{\mathrm{d0}}$ and $\mathrm{d(vtx)_z}/\sigma_{\mathrm{dz}}$ are shown in Figure 5-3. Cutting on the significance of the impact parameter has been preferred with respect the selection on the absolute value of the same quantity. The former distribution have a more uniform behaviour with respect to the latter one. The thresholds have been chosen as a compromise between a high primary particle selection efficiency and low fake rate contamination, as described in Chapter 3.

The first two parameters were adopted in order to remove some of the minimum bias component and to compare with various other experiments such as ALICE and CDF, which have a smaller η range.

Further studies about the tracking quality is given in appendix

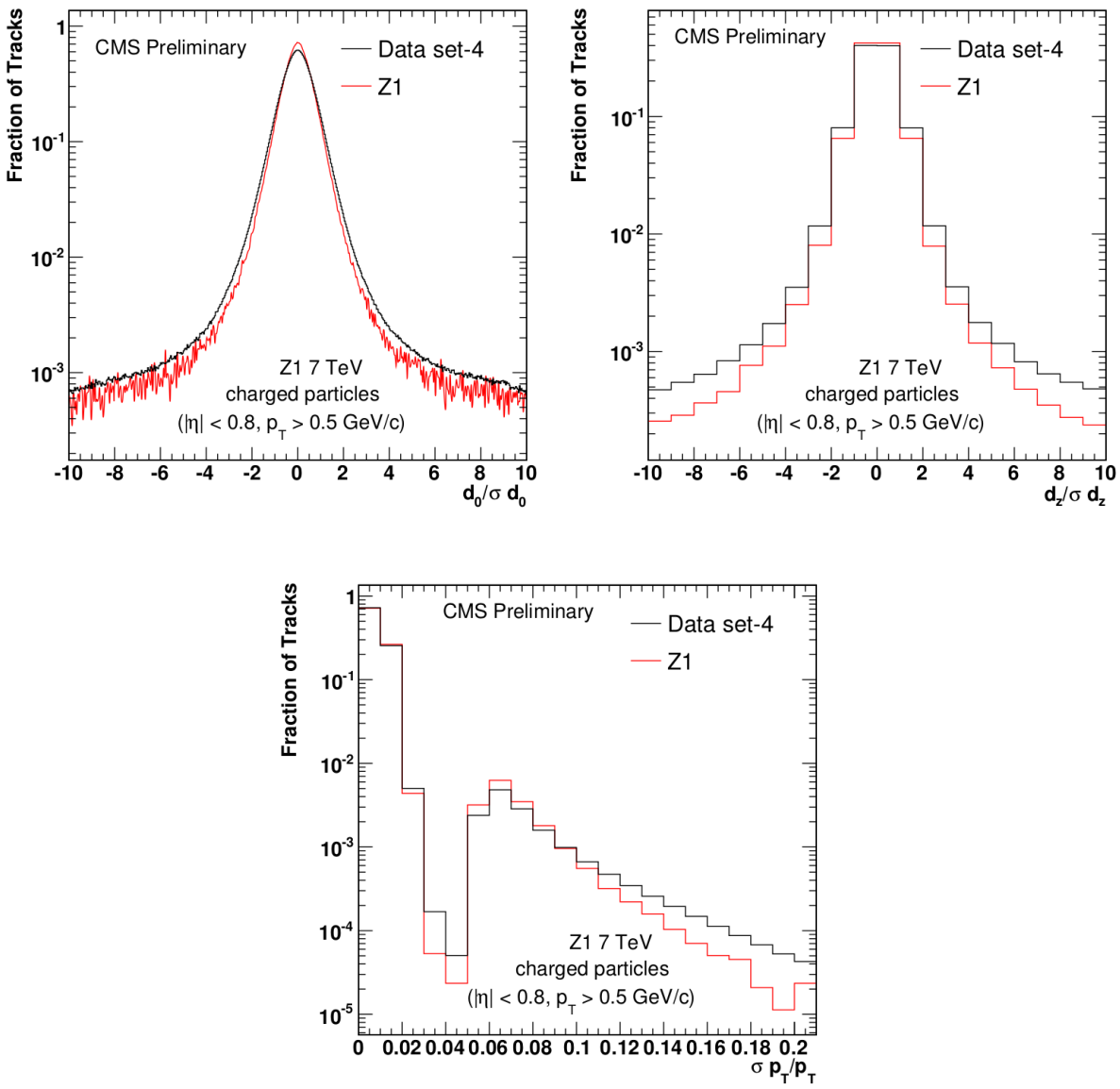

Figure 5-3. $\mathrm{d0}/\sigma(\mathrm{d0})$ (upper left panel), $\mathrm{dz}/\sigma(\mathrm{dz})$ (uper right panel) and the relative $\mathrm{p_T}$ uncertainty at $7\ \mathrm{TeV}$.

5.5 Monte-Carlo Samples

The MC event generator mainly used in this study is Pythia 6.420, it has been the main generatore used in CMS analyses for the first few years of beam usage, there is a slow migration to the newer version Pythia8. As we mentioned in Chapter 2, generators of this sort have a number of free parameters which must be tuned to describe the experimental data accurately. Here we concentrate on the parameters describing the MPI, their evolution as a function of the centre-of-mass energy is described by Equation (2–10). The first successful tune at hadron machines was Tune A developed by the CDF collaboration, primarily by fitting UE data. Starting from this tune quite large number tunes have been developed in the recent years. The main features of the tunes used are:

- Tune D6T was the reference tune for the CMS Collaboration when the LHC started. It was based on CDF run II results and it adopted the CTEQ6L PDFs. It used the virtuality shower and the MPI model developed in PYTHIA 6.2.
- Tune Z1 was the first CMS UE tune. Using Minimum Bias ATLAS tune (AMBTI), R. Field set the PDF to CTEQ5L and fitted PARP(82) and PARP(90) to fit CMS UE data at $7\ \mathrm{TeV}$ and $0.9\ \mathrm{TeV}$.
- Tune Z2* was based on a refit of tune Z2 (a successor of tune D6T, with a refit using CMS UE data in conjunction with CTEQ6L as the PDF) using the results of the most recent underlying event analysis from CMS [27]. The fit was done using PROFESSOR.

Table 5-4. Different parameter values for the 3 used PYTHIA6 tunes.

Parameter	D6T	Z1	Z2*
PDF	CTEQ6L	CTEQ5L	CTEQ6L
MSTP(81)	1	21	21
PARP(82)	1.8387	1.932	1.921
PARP(83)	0.5	0.356	0.356
PARP(84)	0.4	0.651	0.651
PARP(85)	1	0.9	0.9
PARP(86)	1	0.95	0.95
PARP(89)	1.96	1.8	1.8
PARP(90)	0.16	0.275	0.227
PARP(62)	1.25	1.025	1.025
PARP(64)	0.2	1.0	1.0
PARP(67)	2.5	1.0	1.0
PARP(91)	2.1	2.0	2.0
PARP(93)	15.0	10.0	10.0

CHAPTER 6
ANALYSIS RESULTS

The UE is associated with the hadronic activity not included in the hardest 2-2 scattering. The observables used to quantify such activity are the charged particle multiplicity $\mathrm{N_{ch}}$ and the scalar sum of the transverse momenta $\Sigma \mathrm{p_T}$ for the charged particles. These two quantities are usually divided by the appropriate $\eta - \phi$ area of the covered region. Other quantities can be derived from these two variables to test the MC models further.

In order to account for the energy dependence of the UE, these observables are presented at two center of mass energies: $7\ \mathrm{TeV}$ and $0.9\ \mathrm{TeV}$. They are presented as a function of the leading track $\mathrm{p_T}$, which sets the scale for the hard component of the interaction. Predictions from different MC models are presented to compare with the corrected data.

6.1 Hard-scale Dependence at $7\ \mathrm{TeV}$ and $0.9\ \mathrm{TeV}$

This section presents the $\mathrm{N_{ch}}$ and $\Sigma \mathrm{p_T}$ densities at 3 distinct $\eta - \phi$ regions: Toward, Transverse, and Away. These regions are divided with respect to the direction of the leading track, as demonstrated in Figure 6-1.

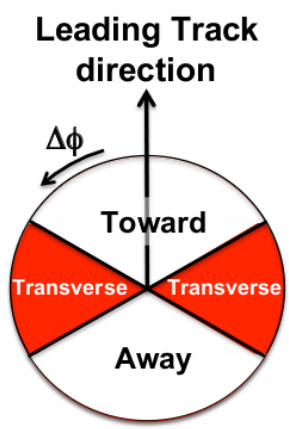

Figure 6-1. The $\eta - \phi$ phase space division scheme with respect to the leading track.

6.2 The UE in the Toward, Away and Transverse Regions

Figure 6-2 shows the average multiplicity in the transverse region divided by the phase space area ($\Delta\eta\Delta(\Delta\phi) = 0.8\text{x}4\pi/3$) as a function of the hardest track. The horizontal error bars indicate the bin size, which tends to increase as we reach higher values of the leading track transverse momentum due to rebinning. There are two error bars at each data point: The inner error bars indicate the amount of statistical uncertainty affecting that point. The outer error bars, which represent the statistical and systematic uncertainty added in quadrature. This convention will apply in all our results, except for the common plots, where the horizontal error bars are omitted for clarity.

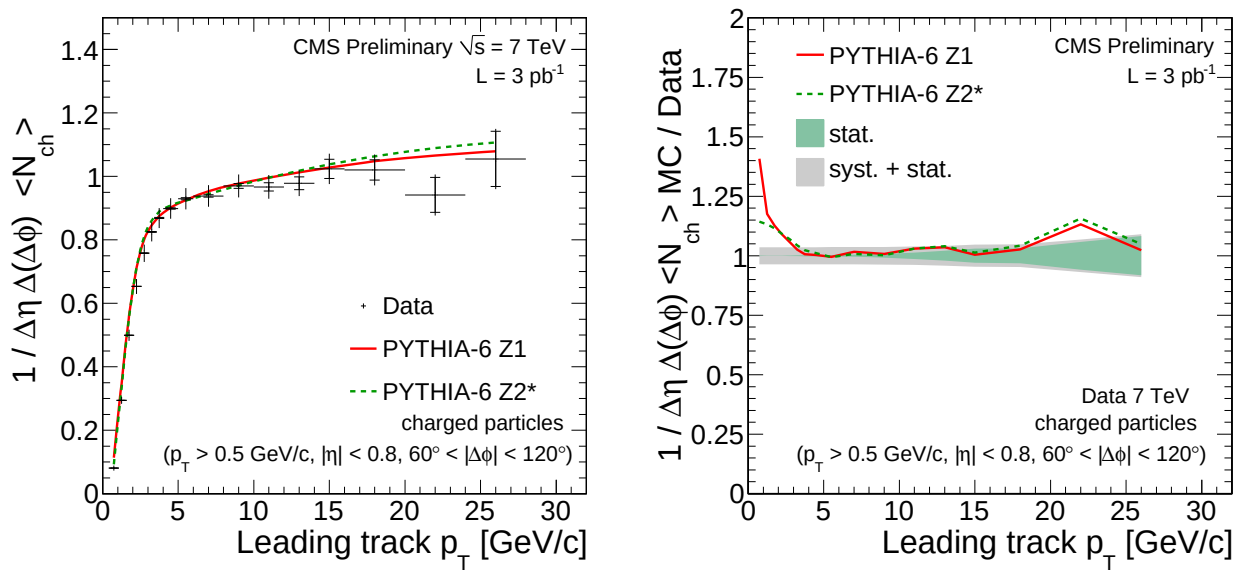

Figure 6-2. The average charge multiplicity as a function of the leading track $\mathrm{p_T}$ in the transverse region at $7\ \mathrm{TeV}$ (right), and the ratio of the MC over Data (left).

In Figure 6-2, two distinctive features can be seen: A rapid rise for regions with $\mathrm{p_T}$ less or equal to 5 GeV/c. This is mainly attributed to the increase of multiple parton interaction activity [46], that corresponds to more centralized collisions (Figure 2-3). The raise is followed by a region of plateau where the multiplicity reaches roughly a constant value that increases slowly. This slight increase is associated with the increase of radiative contributions, the change of energy and fragmentation scale.

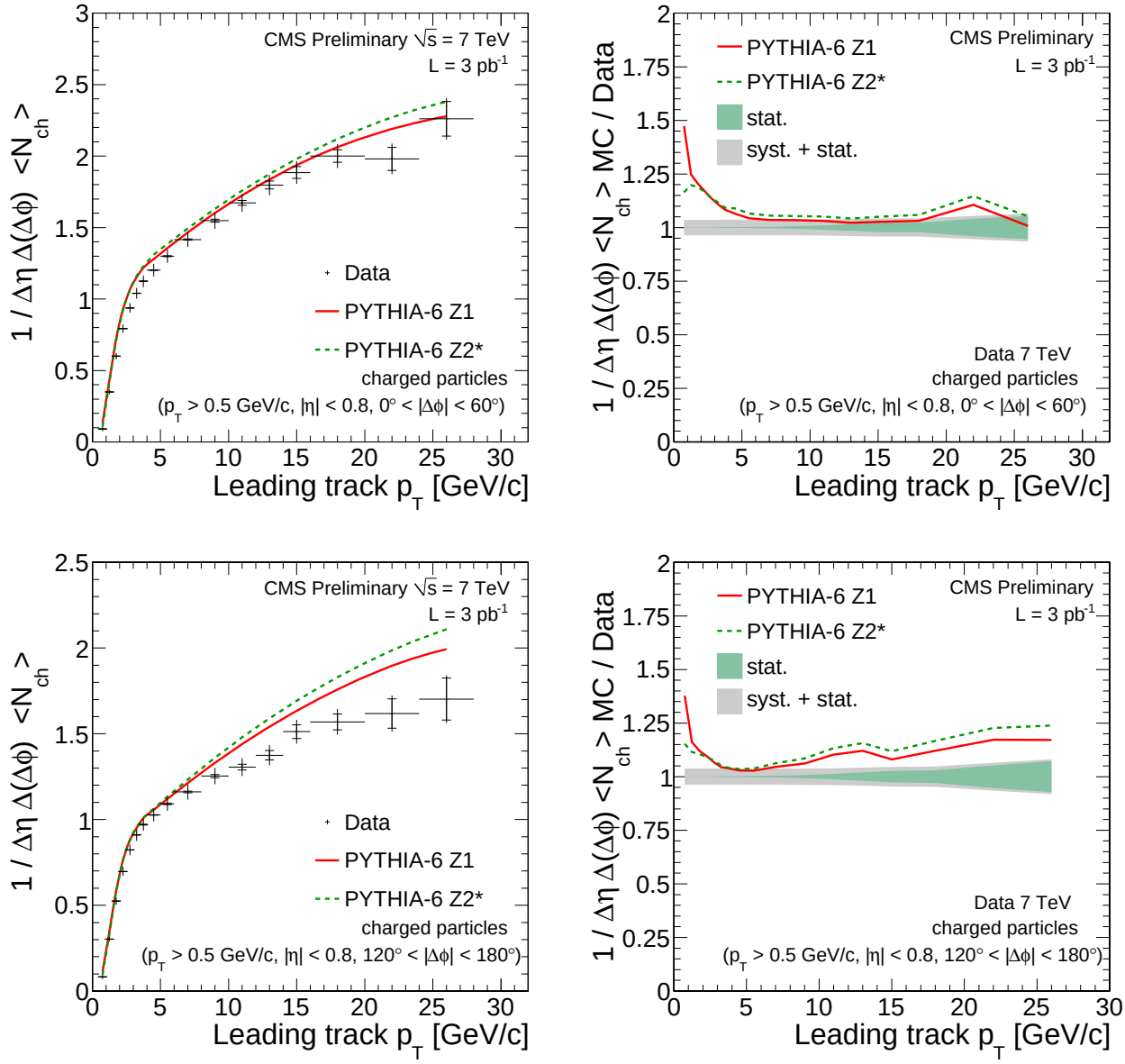

Figure 6-3. The average charge multiplicity as a function of the leading track, for the Toward (upper left panel) and away regions (lower left panel), at $7\ \mathrm{TeV}$. And the ratio of the MC curves over the data for the same $\eta - \phi$ region.

At the right panel of Figure 6-2, the ratio of the MC generators are shown and used to compare with the corrected data. There, the systematic uncertainty is shown in the inner (green) bands and the total uncertainty (again, statistical and systematic uncertainty are added in quadrature) is shown in the gray band. We see that the tunes ($\mathrm{Z1}$ and $\mathrm{Z2^*}$) seem to give a very good qualitative description of the data (right panel). However, when comparing the ratios, it is clear that both tunes seem to overshoot the hadronic activity in the transverse region at the peripheral collisions region. This is mainly due to contribution from diffraction. $\mathrm{Z2^*}$ seems to have lesser deviation from the data at very low values of the leading track momentum. At the other end of the spectrum, the quality of the MC tunes seem to decrease as the statistical uncertainty increases. The transverse region is considered the least complicated among the other regions under consideration in this study. The hadronic activity is sensitive to BBR and to MPI.

A similar pattern is observed for the Toward and Away $\mathrm{N_{ch}}$ density (Figure 6-3); however, it is clear that the activity at the centralized region increase at a higher rate with respect to $\mathrm{p_T}$ of the leading track compared to the Transverse region. This is due to ISR and FSR contribution from the jets in the Toward and Away regions. Both tunes give a good qualitative description of the data; however, they predict much higher multiplicity at the low values of the leading track $\mathrm{p_T}$. At the other extreme, they also give a reduced accuracy (much worse for the Away region where the tunes seem to overestimate the hadronic activity) due to statistical fluctuations.

The Toward is a complicated region to model. It is sensitive to fragmentation and it receives contributions from BBR, MPI, and ISR. The away region is considered even more complicated. It depends if the away-jet lies outside the $|\eta| < 0.8$ cut or not. If the away-jet is outside the cut, the region recieves contributions from the away-jet, BBR , MPI, ISR, FSR, and PDF. The last factor emerges from the contribution of the 2-2

subprocesses involved (the type of partons involved). This level of complication makes it difficult for the MC models to simulate the data accurately.

The second observable used to study the hadronic activity is the scalar sum for the Transverse momentum $\Sigma \mathrm{p_T}$. Figures 6-4, and 6-5 show the behavior of this observable as a function of the leading track Transverse momentum. $\Sigma \mathrm{p_T}$ plots have the same general features as $\mathrm{N_{ch}}$ for the three phase space regions. The similarity of the MC predictions is visible; i.e. good qualitative agreement, and poor agreement at the qualitative scale at the peripheral region due to diffraction contamination. We also observe some deviation at the high end of the spectrum that contains lower statistics, it is observed that the MC predictions are much worse at the Away region.

Similar studies are performed at $0.9\ \mathrm{TeV}$. Figures 6-6 and 6-7 show the charge multiplicity behaviour with respect to the leading track. The qualitative features are well represented by the MC models. The qualitative features are improved and are much closer to the data for the peripheral region for tune $\mathrm{Z2^*}$, while tune $\mathrm{Z1}$ shows a similar divergence to the one shown at $7\ \mathrm{TeV}$. The same can be said for the Toward and Away regions, with the Away region showing a higher activity for the charge multiplicity for both tunes.

For $\Sigma \mathrm{p_T}$ at $0.9\ \mathrm{TeV}$, tune $\mathrm{Z2^*}$ gives a much better quantitative description of the data at the soft collisions region, while $\mathrm{Z1}$ shows divergence from data at $\mathrm{p_T}$ for the leading track gets lower.

6.3 The Maximum and Minimum Transverse Regions

There are further constrains that can be imposed on the $\eta - \phi$ space. For instance, the Transverse region can be divided into two equal regions: one of them containing more hadronic activity called the maximum (max) region, and the other region will be called the minimum (min) region. The max region has a higher probability of containing a third jet. Both regions receive contributions from multiple parton interactions and beam beam remnants. The max region sensitive to many factors just like the Toward and Away

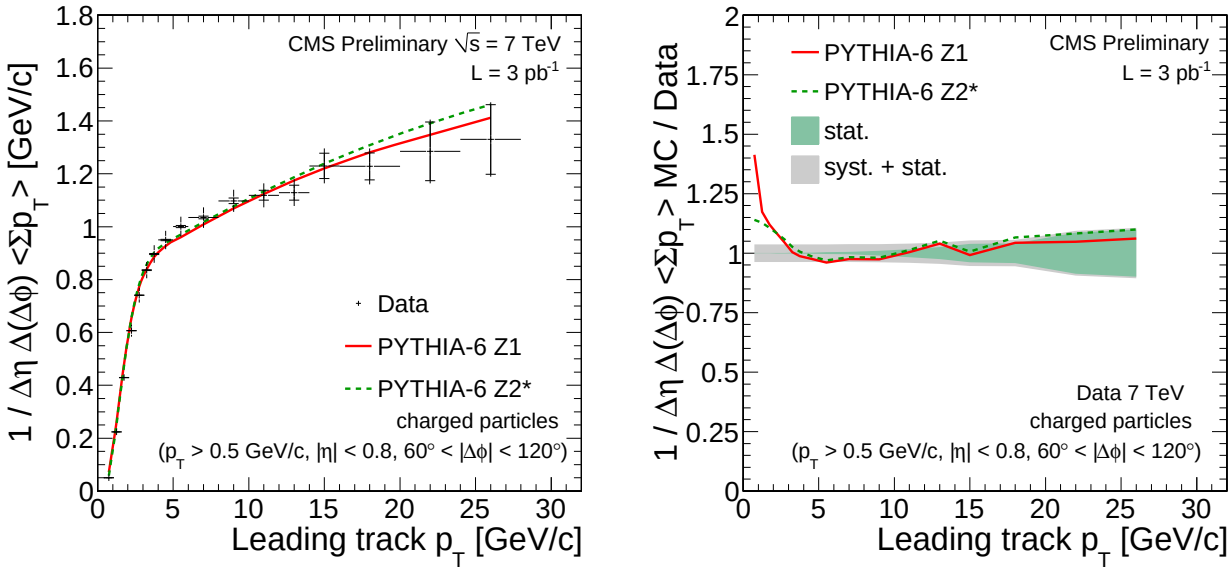

Figure 6-4. The average Transverse momentum sum as a function of the leading track $\mathrm{p_T}$ in the Transverse region at $7\ \mathrm{TeV}$ (right), and the ratio of the MC over Data (left).

regions, while the min region is more sensitive to MPI and BBR in particular. We can also study the difference between the two regions for both observables (dif), which is more sensitive to ISR and FSR. The difference between the two regions is sensitive to the final state radiation. Figures 6-10 and 6-12 show the average multiplicity profile plots at the max ($\mathrm{Nch_{Max}}$ and $\Sigma\mathrm{p_{TMax}}$) and min ($\mathrm{Nch_{Min}}$ and $\Sigma\mathrm{p_{TMin}}$) regions. We see an excellent agreement for the $\mathrm{Nch_{Max}}$ throughout the spectrum, except for the first few bins. For $\mathrm{Nch_{Min}}$, we notice a good agreement between the MC and the data, except for the two extremes of the spectrum. The agreement is better for $\mathrm{Nch_{Max}}$ compared to $\mathrm{Nch_{Min}}$. A similar pattern is observed for $\Sigma\mathrm{p_{Tmax}}$ and $\Sigma\mathrm{p_{Tmin}}$.

The same sector was studied for $0.9\ \mathrm{TeV}$ center of mass energy. There, a good agreement is also noticed (not as good as the case of $\mathrm{Nch_{Max}}$ at $7\ \mathrm{TeV}$) for both observables, with the biggest discrepancy being reported at the lowest and highest bins. See Figures 6-14 and 6-16.

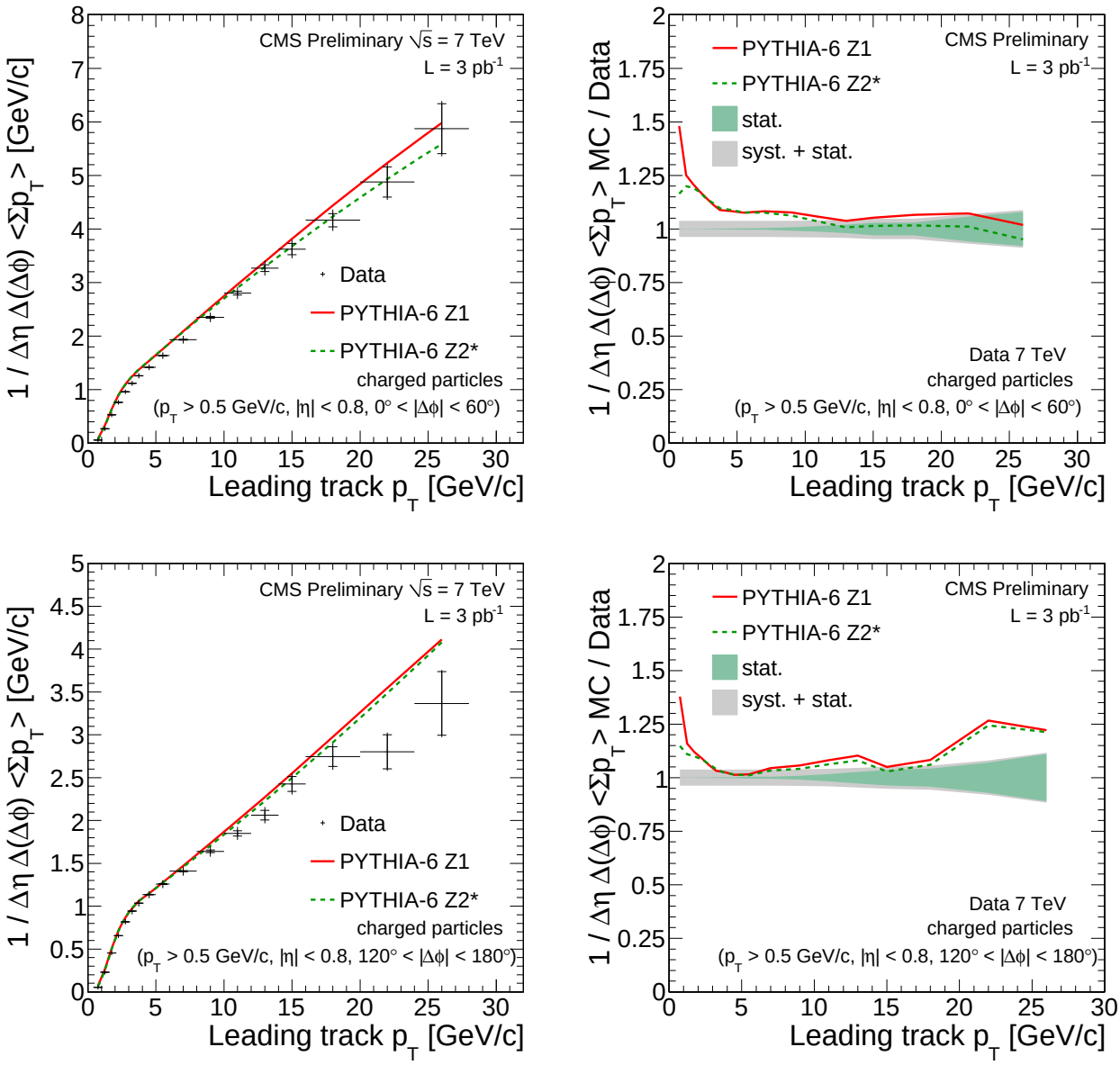

Figure 6-5. The average Transverse momentum as a function of the leading track, for the Toward (upper left panel) and Away regions (lower left panel), at TeV. The ratio of the MC curves over the data for the same $\eta - \phi$ regions are shown in the right panel.

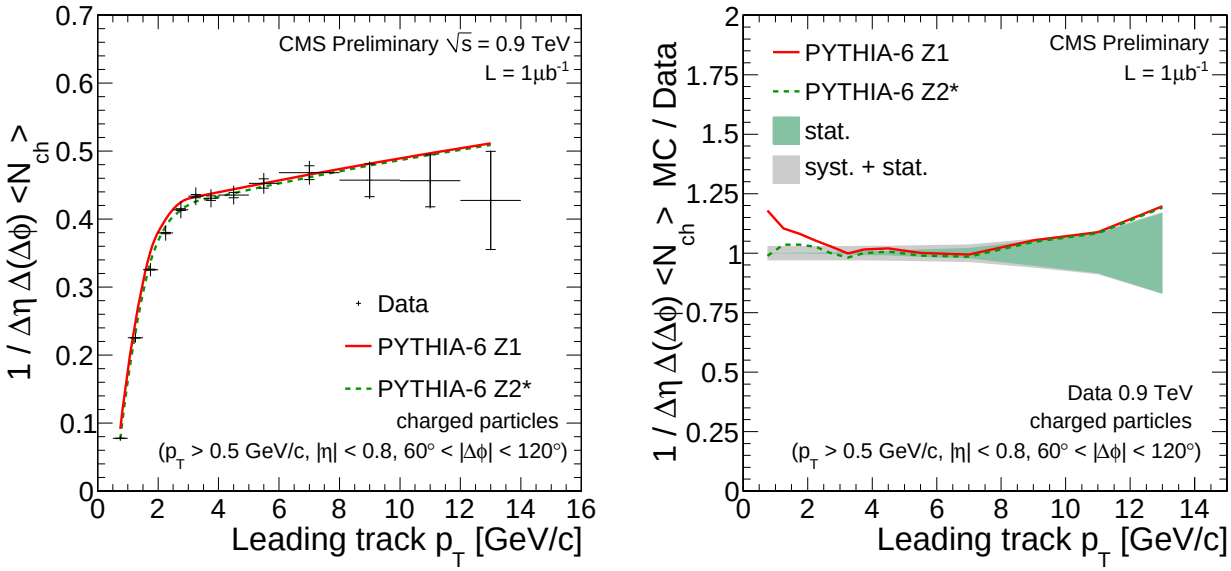

Figure 6-6. The average charge multiplicity as a function of the leading track pT in the Transverse region at $0.9\ \mathrm{TeV}$ (right), and the ratio of the MC over Data (left).

Figures 6-11 and 6-13 shows the difference of the UE activity between the Max and Min regions for Nch and $\Sigma \mathrm{p_T}$ at $7\ \mathrm{TeV}$. There is an excellent overall agreement between the two MC tunes and the data, with the exception of the first few bins. The quality of the MC description deteriorates for both observables at $0.9\ \mathrm{TeV}$, as can be seen in Figures 6-15 and 6-17.

6.4 Average $\mathrm{p_T}$

The next set of observables are constructed from taking the ratios of the observables measured previously at the Transverse region: The bin by bin ratio of $\Sigma \mathrm{p_T}$ over $\mathrm{N_{ch}}$. This observable relies mostly on the shape of the $\mathrm{p_T}$ spectrum emerging from MPI, this links it to color connection schemes.

Figure 6-18 shows the AvepT plots at $7\ \mathrm{TeV}$. Here we see the two tunes giving an almost identical value of the activity throughout the spectrum. When compared to data, they both tend to undershoot the value predicted by the data at all regions except the lowest $\mathrm{p_T}$ values of the leading track. For $0.9\ \mathrm{TeV}$ (Figure 6-19) the two MC tunes show identical predictions at low $\mathrm{p_T}$. However, tune $\mathrm{Z2^*}$ starts showing slightly higher values

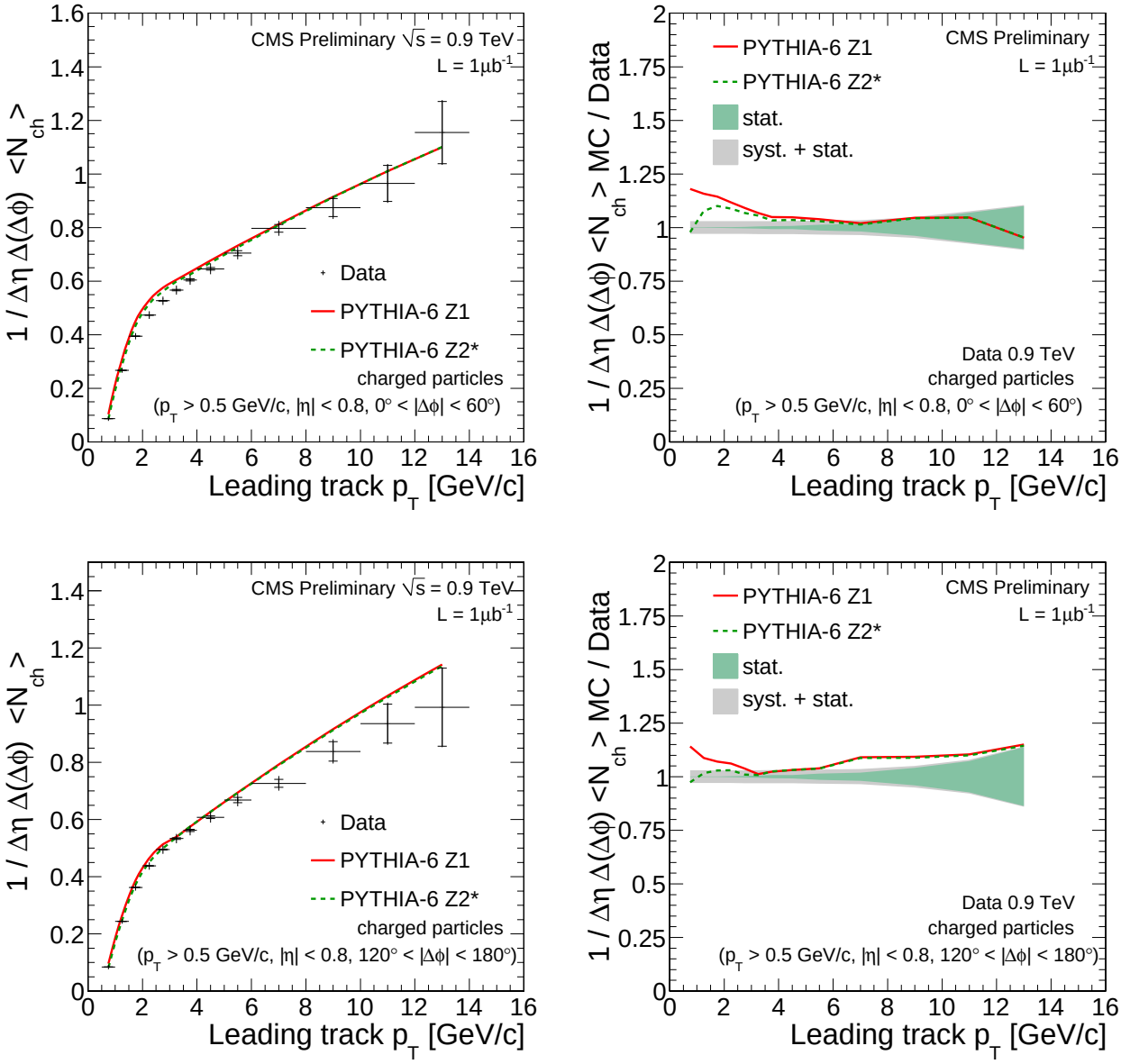

Figure 6-7. The average charge multiplicity as a function of the leading track, for the Toward (upper left panel) and Away regions (lower left panel), at 0.9 TeV. And the ratio of the MC curves over the data for the same $\eta - \phi$ regions.

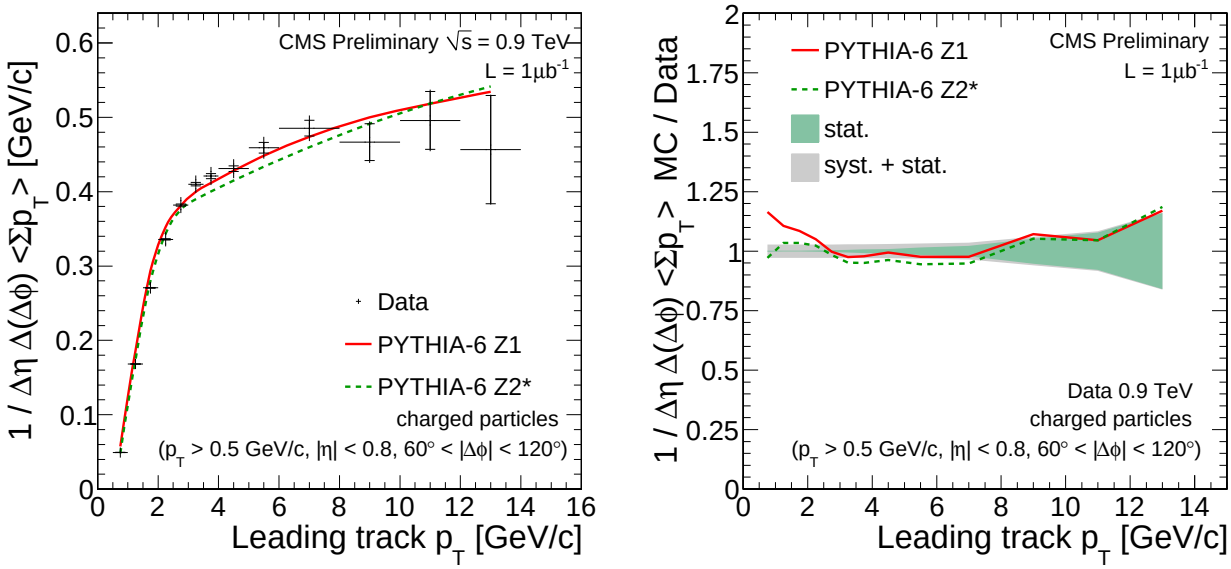

Figure 6-8. The average Transverse momentum sum as a function of the leading track pT in the Transverse region at $0.9\ \mathrm{TeV}$ (right), and the ratio of the MC over Data (left).

at the high end of the spectrum. Both tunes, however, show a better agreement with the data than in $7\ \mathrm{TeV}$.

6.5 The Ratio of the Observables at Two Center-of-Mass Energies

The next set of plots shows the average charge multiplicity and $\Sigma \mathrm{p_T}$ for the two energies studied. (as shown in Figure 6-20). Here we see the two tunes tend to agree with the data behavior. Albeit there is a large divergence at the first bin and it shows less activity at the higher end of the spectrum.

6.6 Comparison with Other Experiments

The results of the UE analysis are compared with the findings of ALICE [47], one of the other major experiments at the LHC. Figure 6-21 shows the similarity of the findings of both collaborations at the two center-of-mass energy levels. CMS results extend to larger values of $\mathrm{p_{Tmax}}$ due to better statistics.

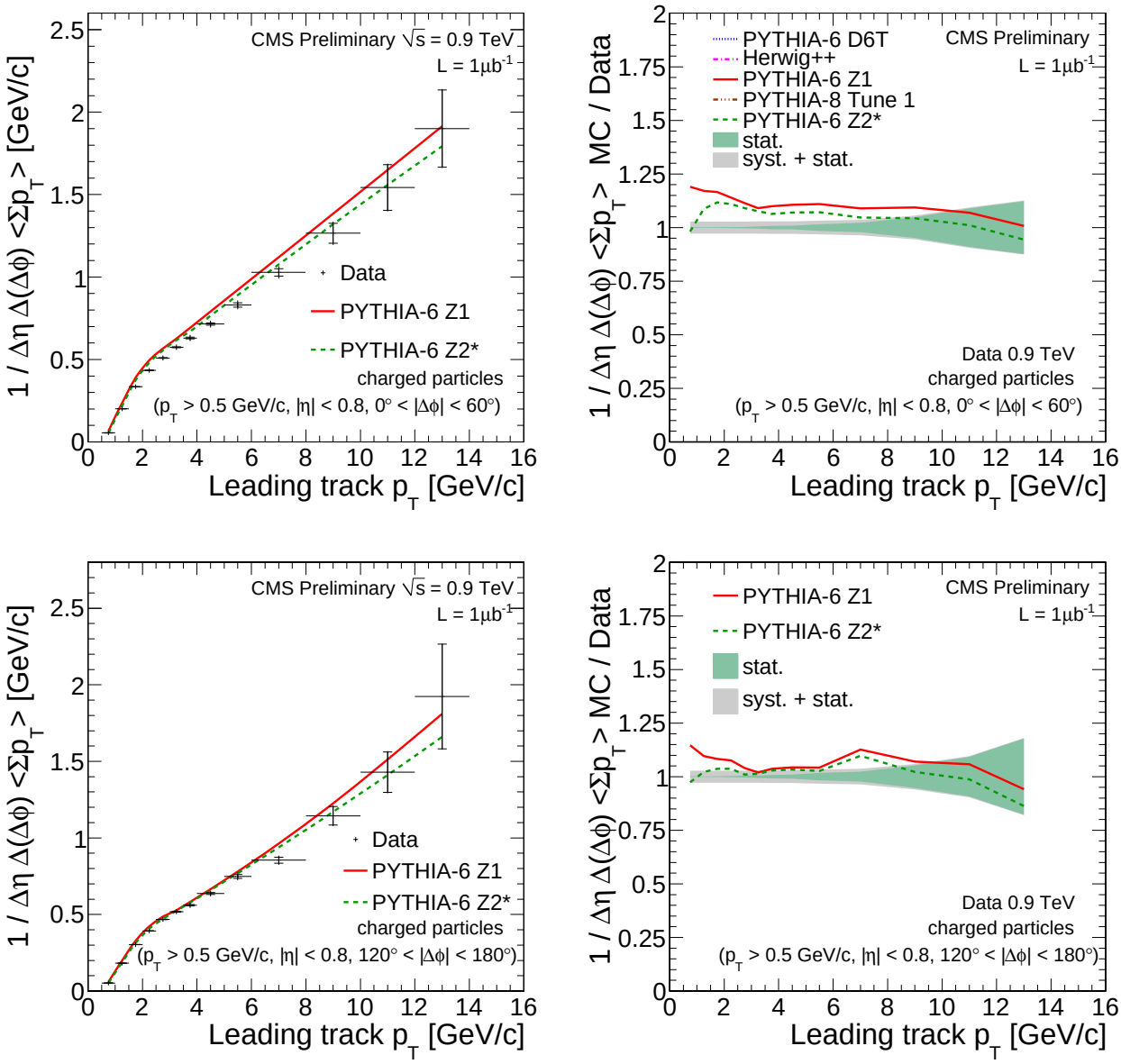

Figure 6-9. The average scalar momentum as a function of the leading track, for the Toward (upper left panel) and Away regions (lower left panel), at 0.9 TeV. And the ratio of the MC curves over the data for the same eta-phi regions.

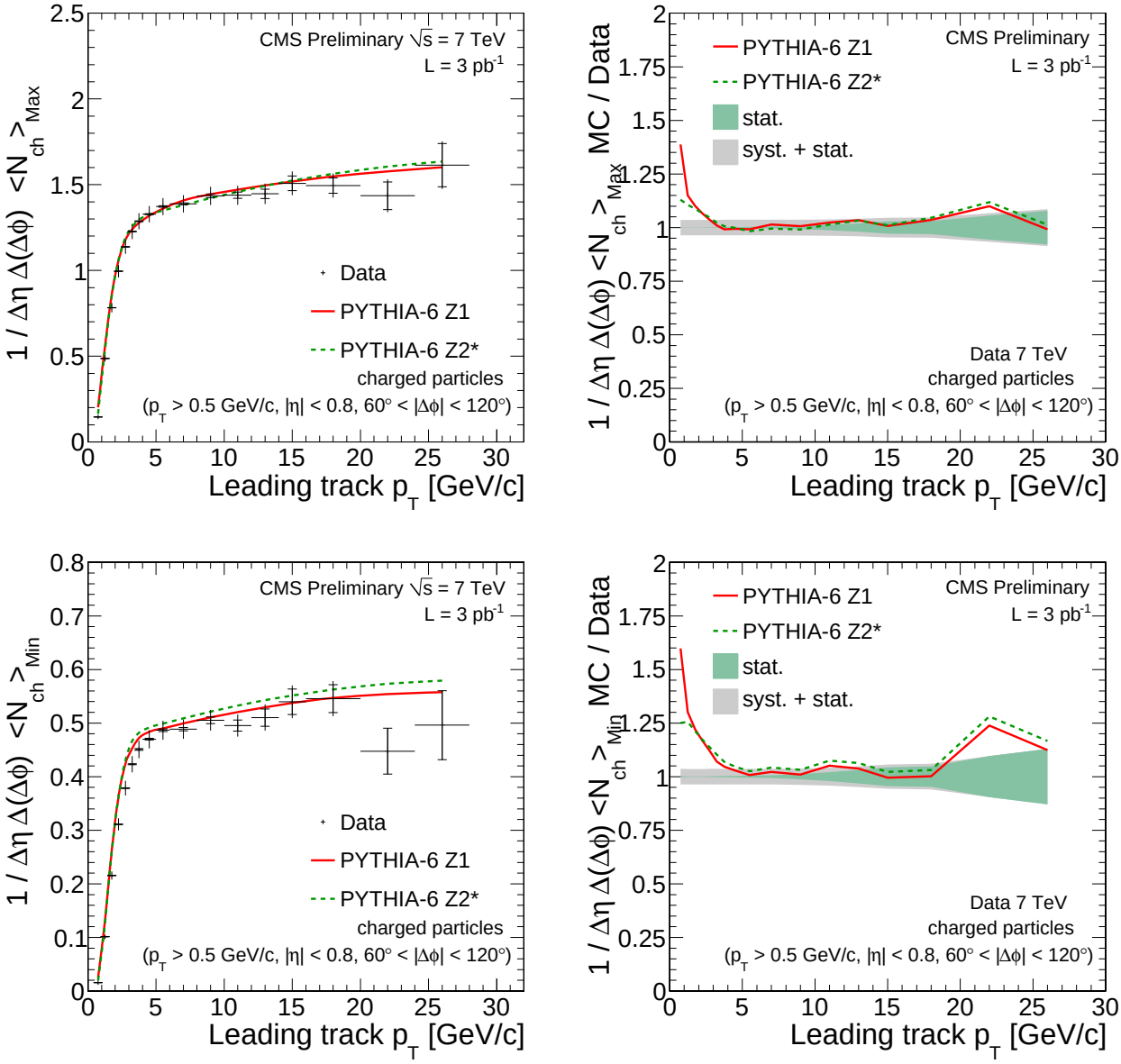

Figure 6-10. The average charge multiplicity as a function of the leading track pT in the maximum (upper left) and minimum (lower left) Transverse region at $7\ \mathrm{TeV}$, and the ratio of the MC over Data (right).

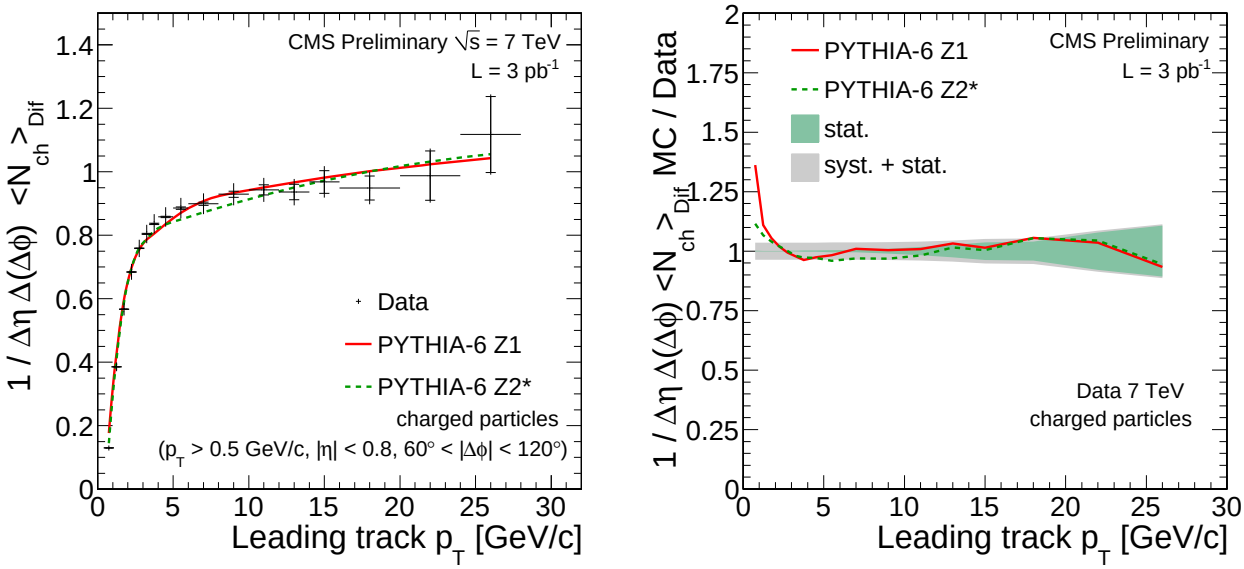

Figure 6-11. The difference in the average charge multiplicity as a function of the leading track pT between the maximum and minimum Transverse regions at $7\ \mathrm{TeV}$ (left), and the ratio of the MC over Data (right).

6.7 Energy Scan of the UE

Taking similar measurements for the hadronic activity at the phase space regions mentioned above at different center-of-mass energies gives a great opportunity to study the energy dependence of the various factors contributing to the UE. This can be seen from Figures 6-22 and 6-23. This energy scan is obtained by combining the results obtained above from CMS with the results obtain by CDF at two center-of-mass energies that have not been explored by the LHC: $0.3\ \mathrm{TeV}$ and $1.96\ \mathrm{TeV}$ [48].

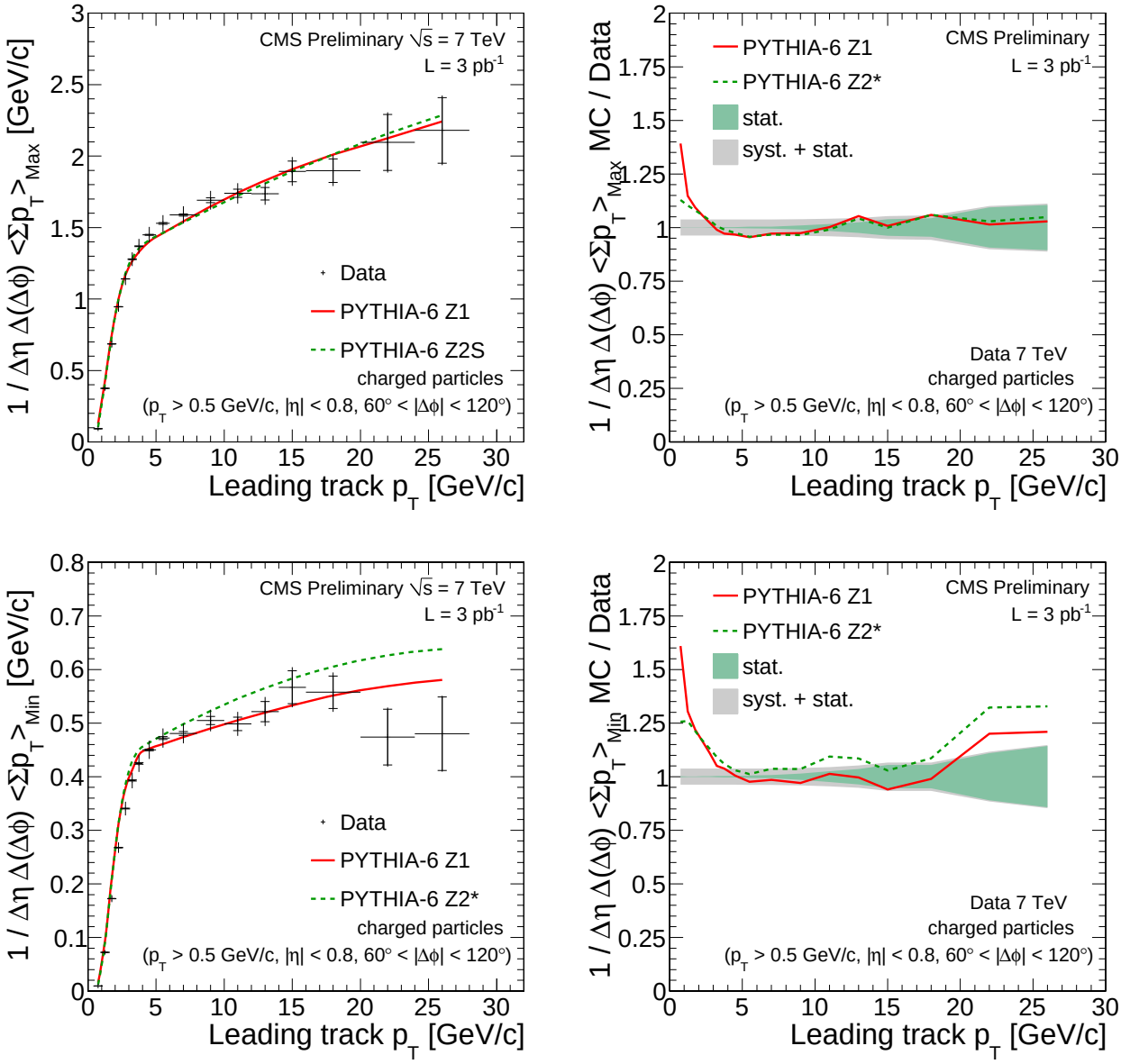

Figure 6-12. The average scalar momentum sum as a function of the leading track $\mathrm{p_T}$ in the maximum (upper left) and minimum (lower left) Transverse region at $7\ \mathrm{TeV}$, and the ratio of the MC over Data (right).

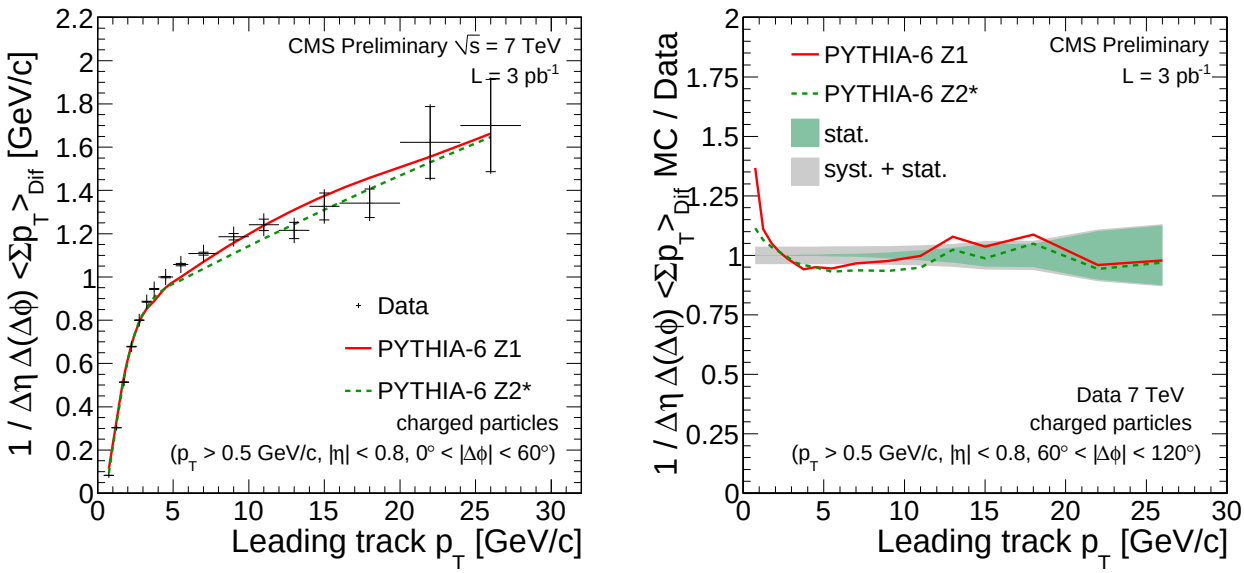

Figure 6-13. The difference in the average scalar momentum sum as a function of the leading track $\mathrm{p_T}$ between the maximum and minimum Transverse regions at $7\ \mathrm{TeV}$ (left), and the ratio of the MC over Data (right).

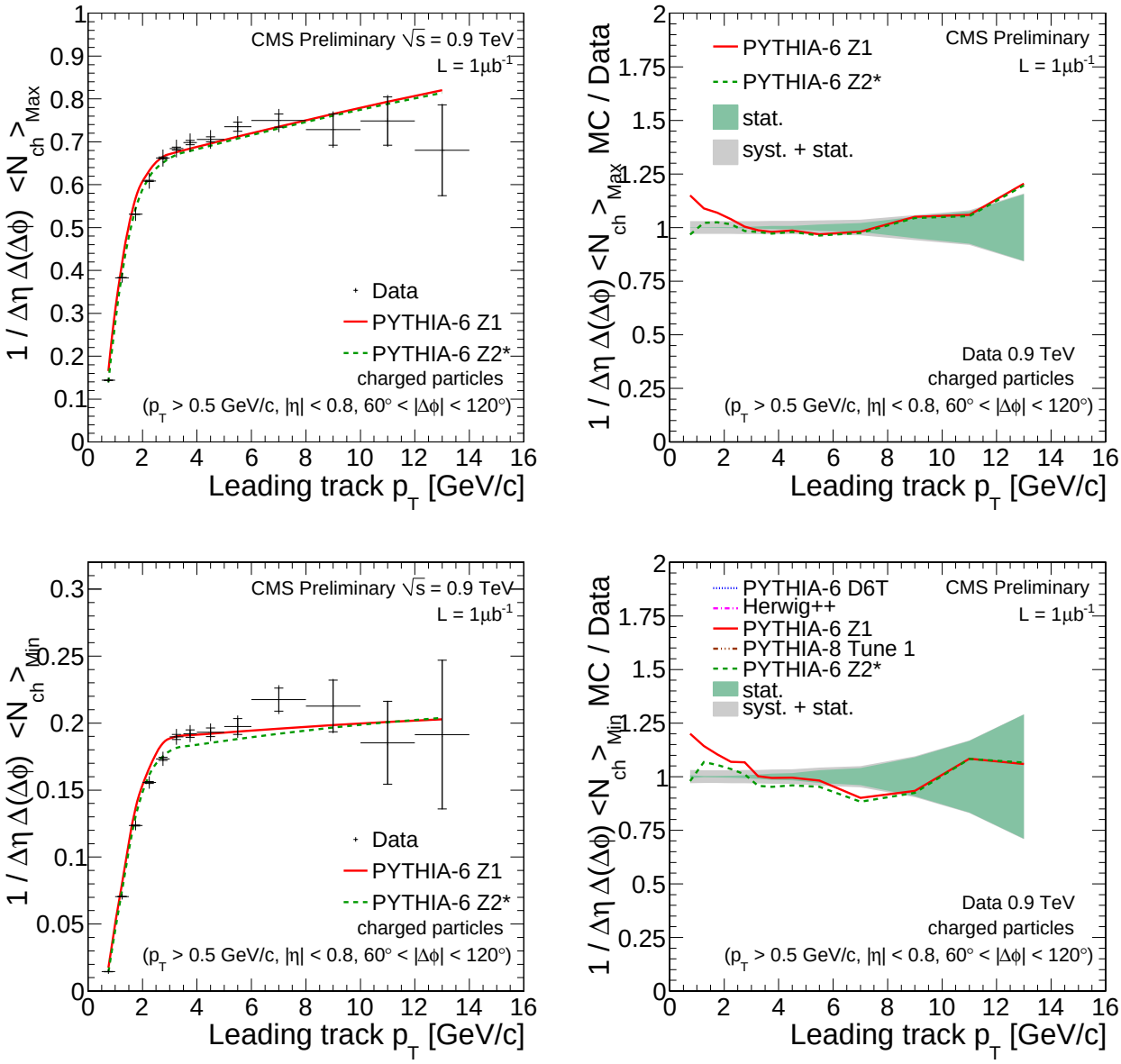

Figure 6-14. The average charge multiplicity as a function of the leading track $\mathrm{p_T}$ in the maximum (upper left) and minimum (lower left) Transverse region at $0.9\ \mathrm{TeV}$, and the ratio of the MC over Data (right).

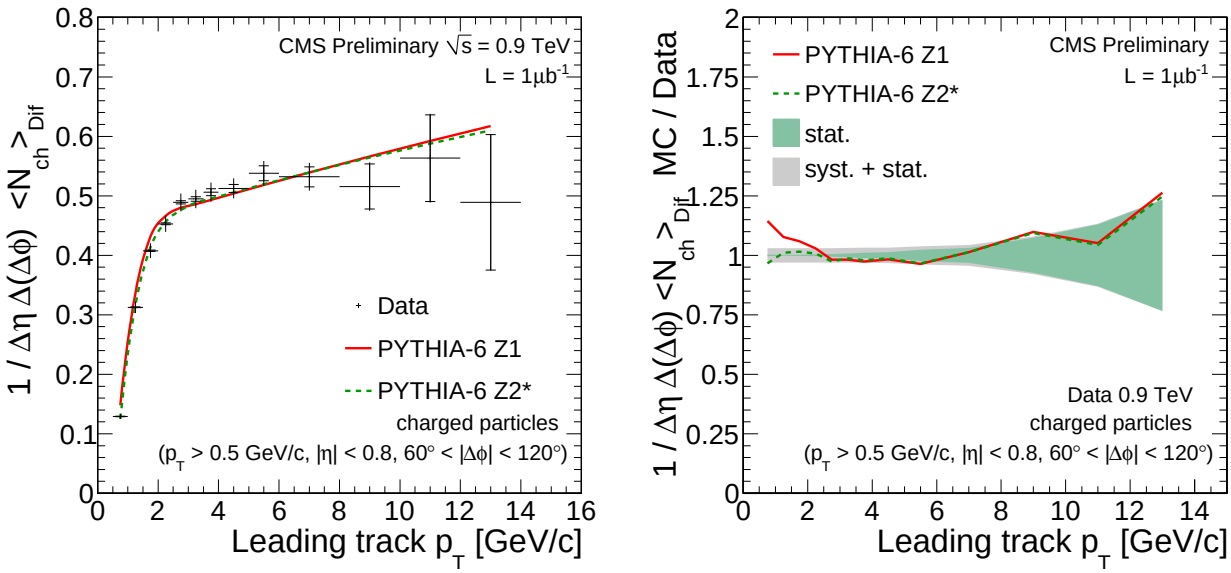

Figure 6-15. The difference in the average charge multiplicity as a function of the leading track $\mathrm{p_T}$ between the maximum and minimum Transverse regions at $0.9\ \mathrm{TeV}$ (left), and the ratio of the MC over Data (right).

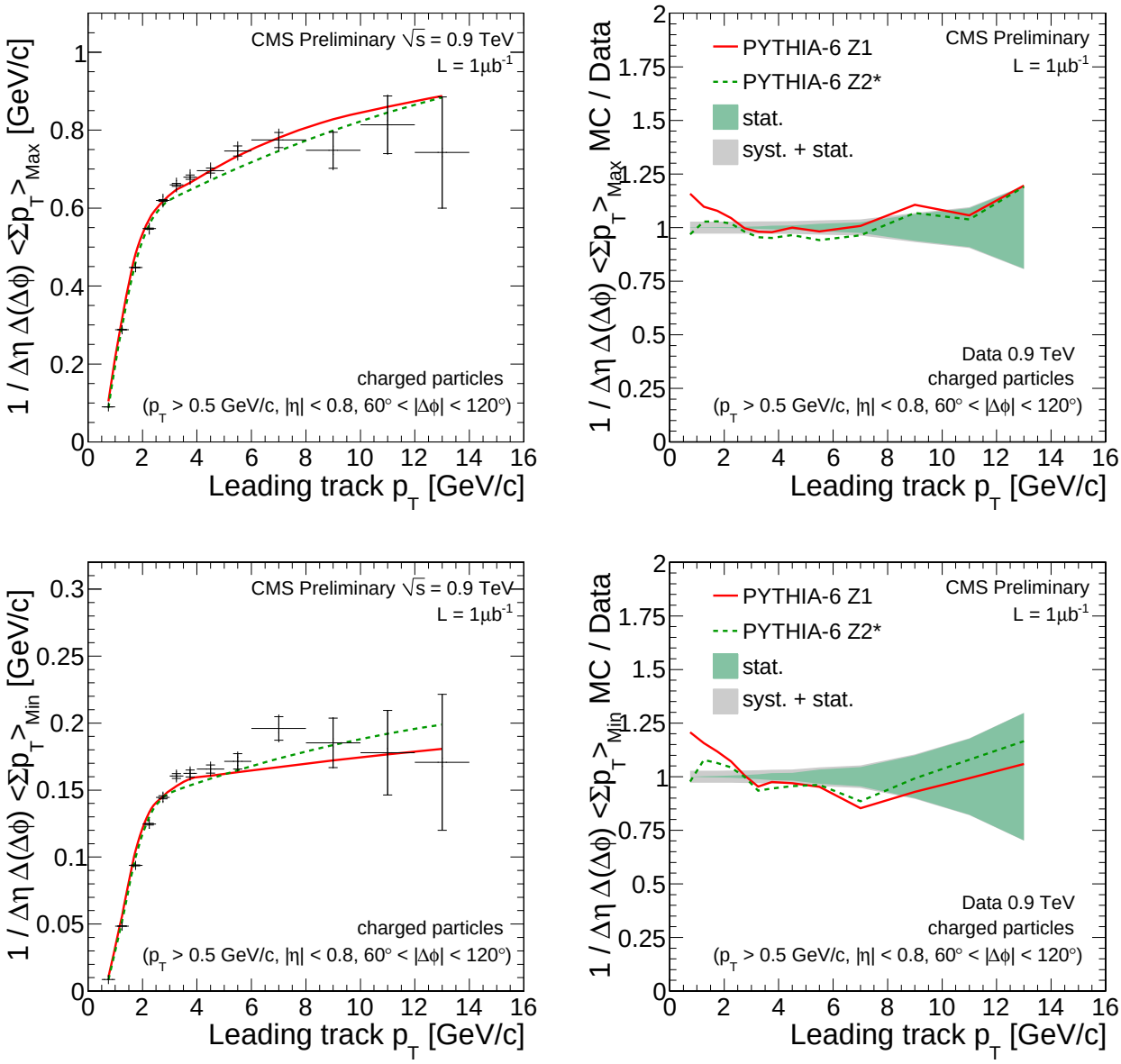

Figure 6-16. The average scalar momentum sum as a function of the leading track $\mathrm{p_T}$ in the maximum (upper left) and minimum (lower left) Transverse region at $0.9\ \mathrm{TeV}$, and the ratio of the MC over Data (right).

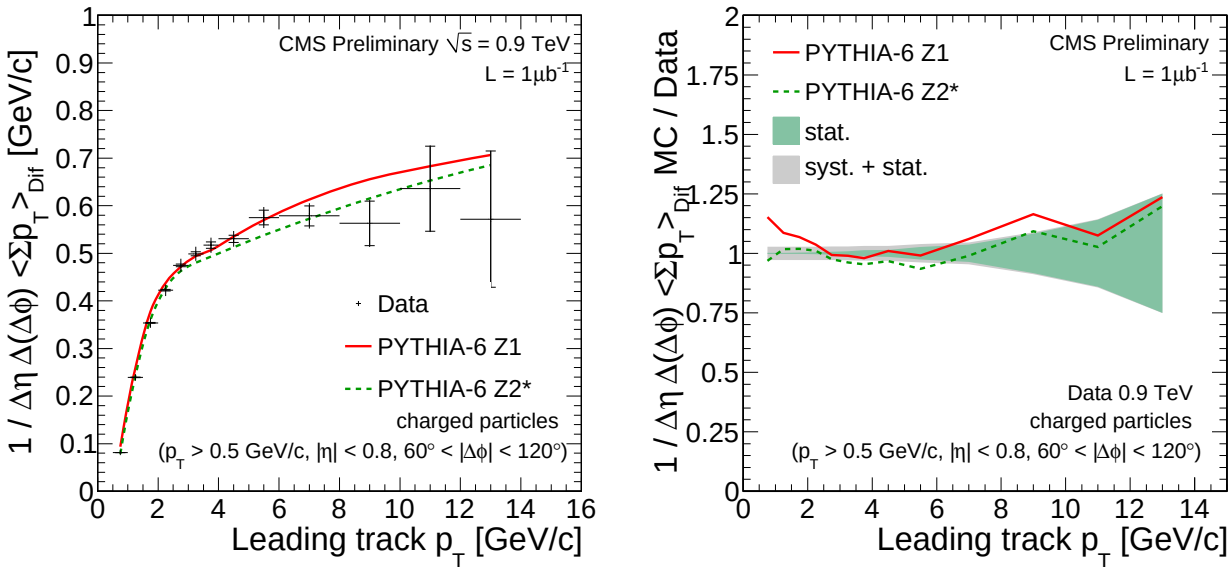

Figure 6-17. The difference in the average scalar momentum sum as a function of the leading track $\mathrm{p_T}$ between the maximum and minimum Transverse regions at $0.9\ \mathrm{TeV}$ (left), and the ratio of the MC over Data (right).

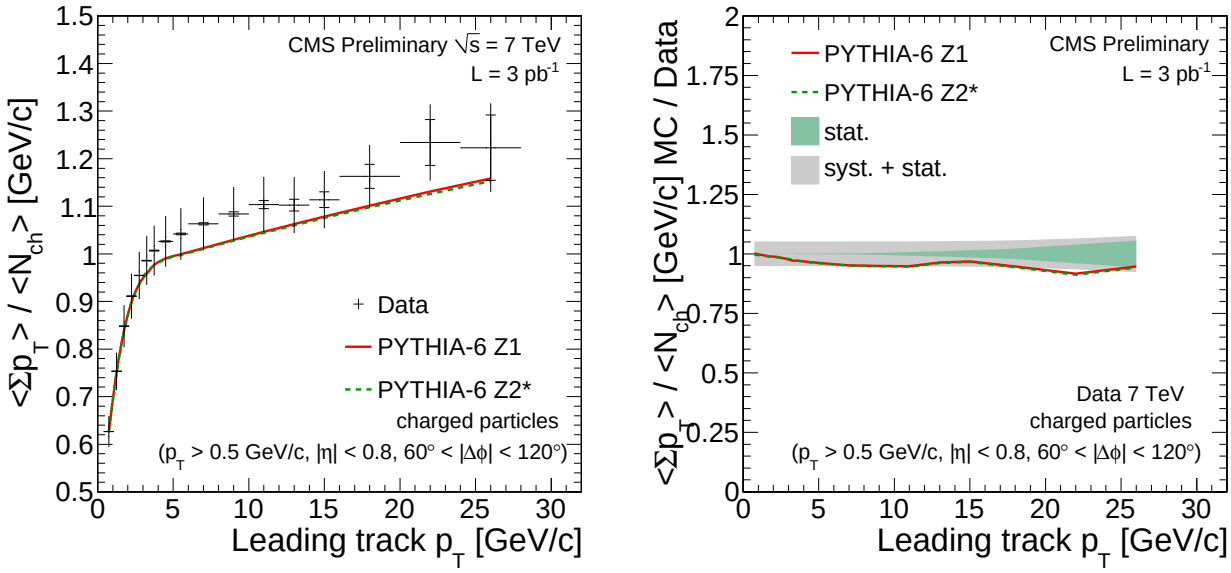

Figure 6-18. The average Transverse momentum as a function of the leading track $\mathrm{p_T}$ at $7\ \mathrm{TeV}$ in the Transverse region.

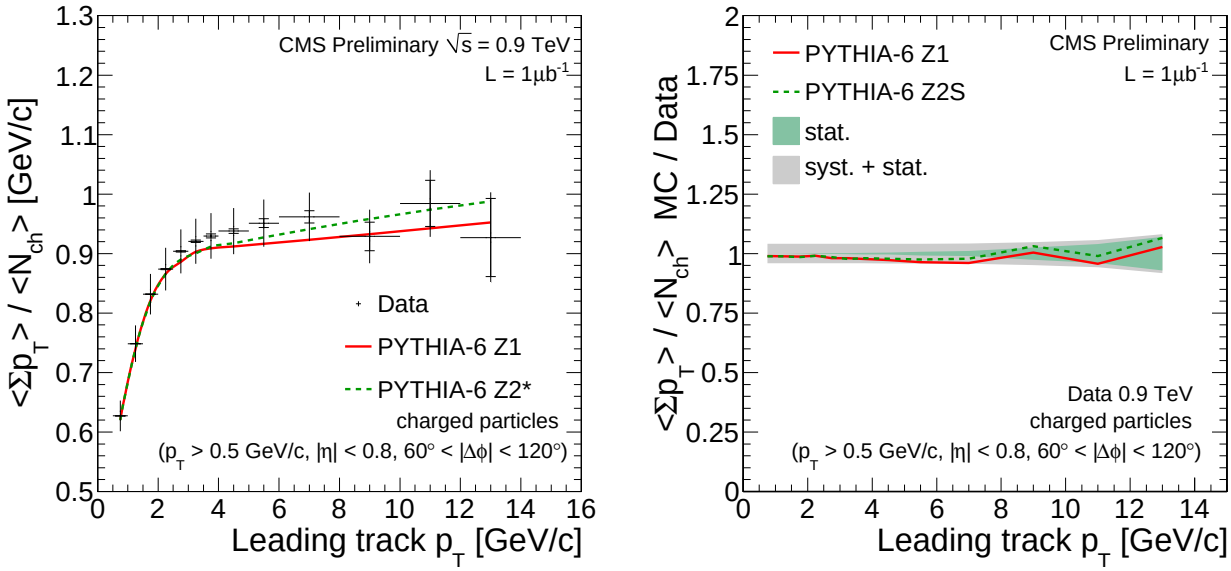

Figure 6-19. The average Transverse momentum as a function of the leading track $\mathrm{p_T}$ at $0.9\ \mathrm{TeV}$ in the Transverse region.

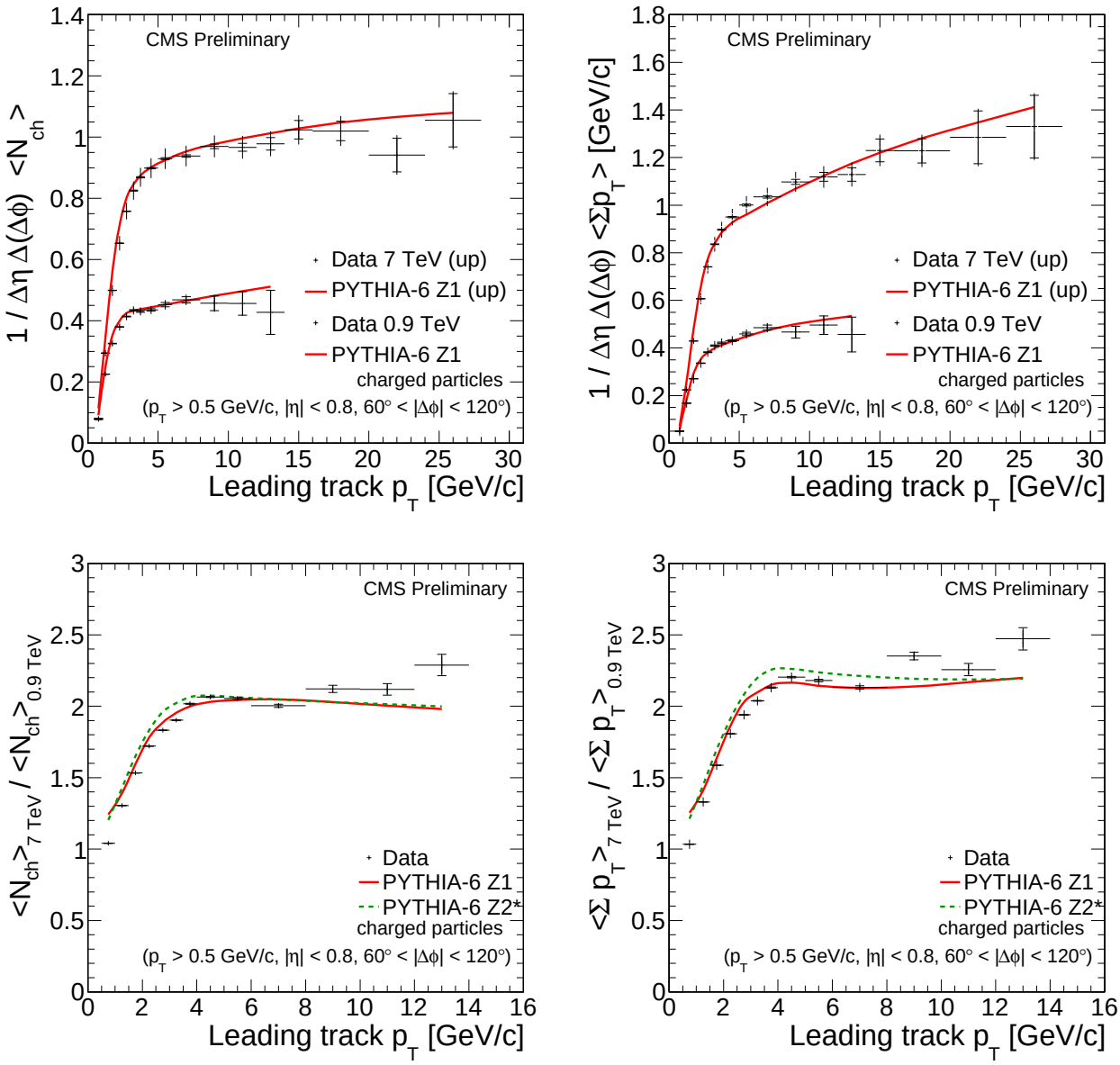

Figure 6-20. The average charge multiplicity (upper left) and scalar momentum sum (upper right) as a function of the leading track pT in the Transverse region at $7\ \mathrm{TeV}$ and $0.9\ \mathrm{TeV}$. The ratio of the same observables at the two center of mass energies.

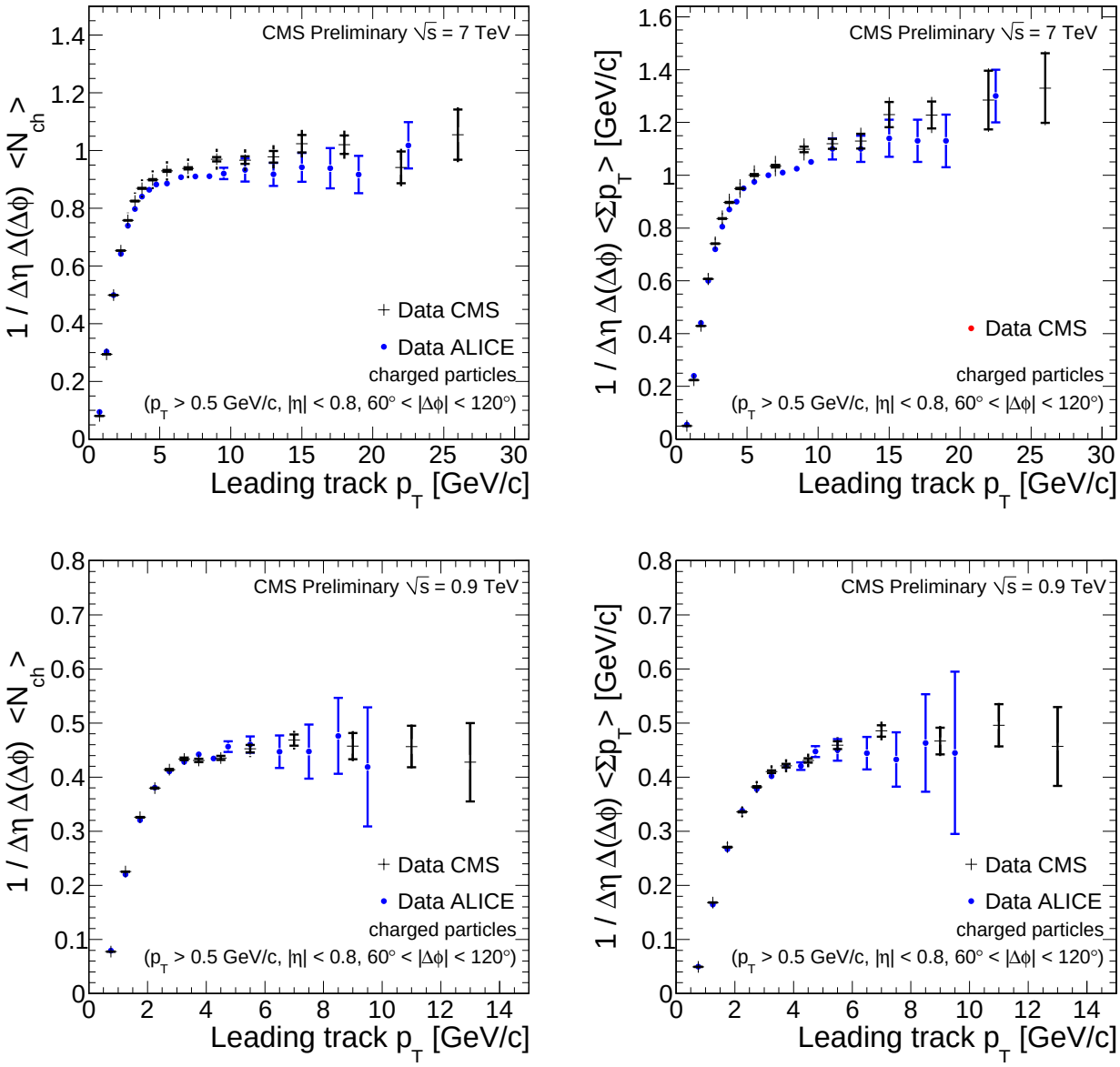

Figure 6-21. The average charge multiplicity and the Transverse momentum sum as a function of the leading track for CMS and ALICE.

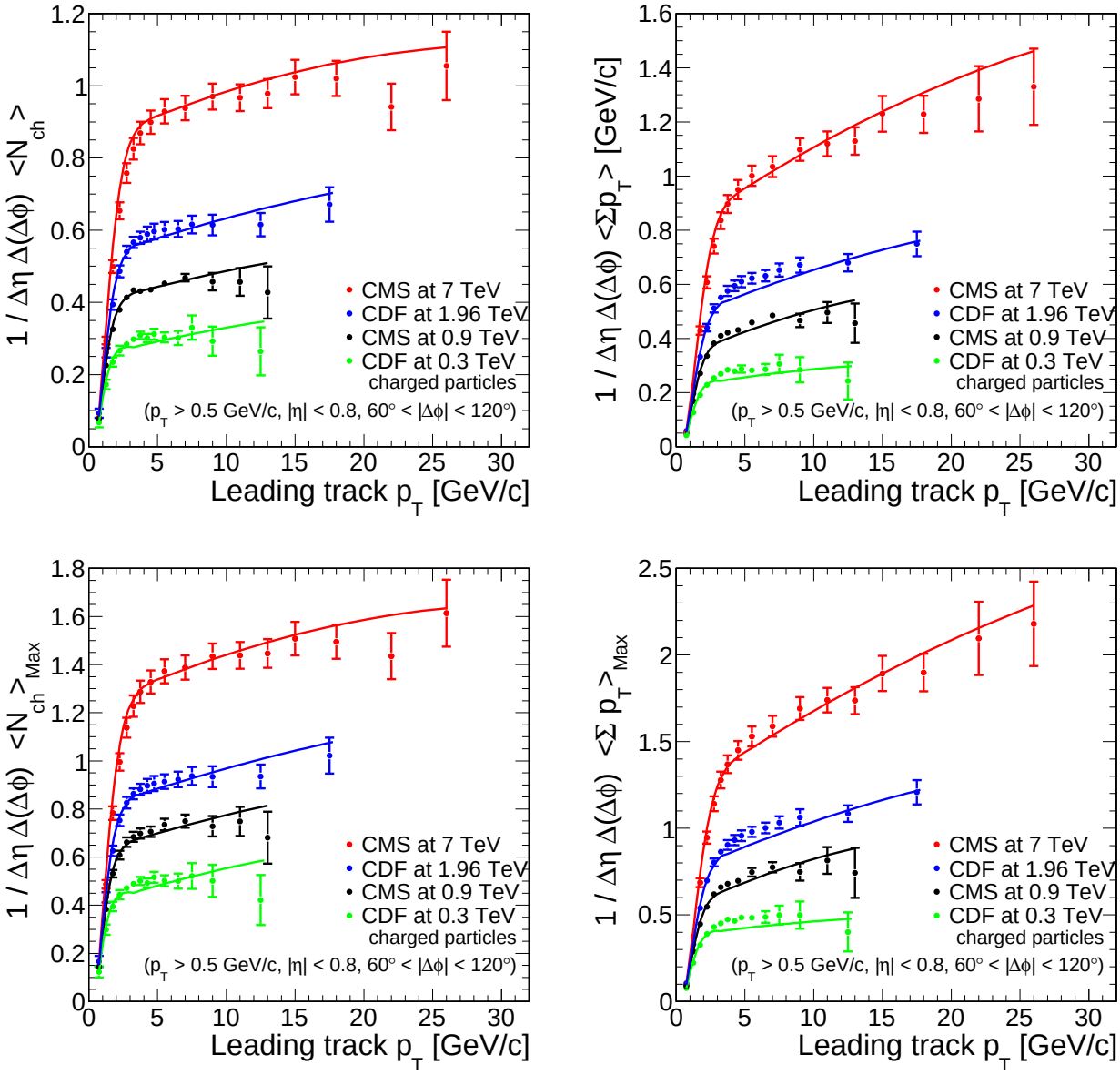

Figure 6-22. The hadronic activity at four center-of-mass energies in the Transverse region and the max region.

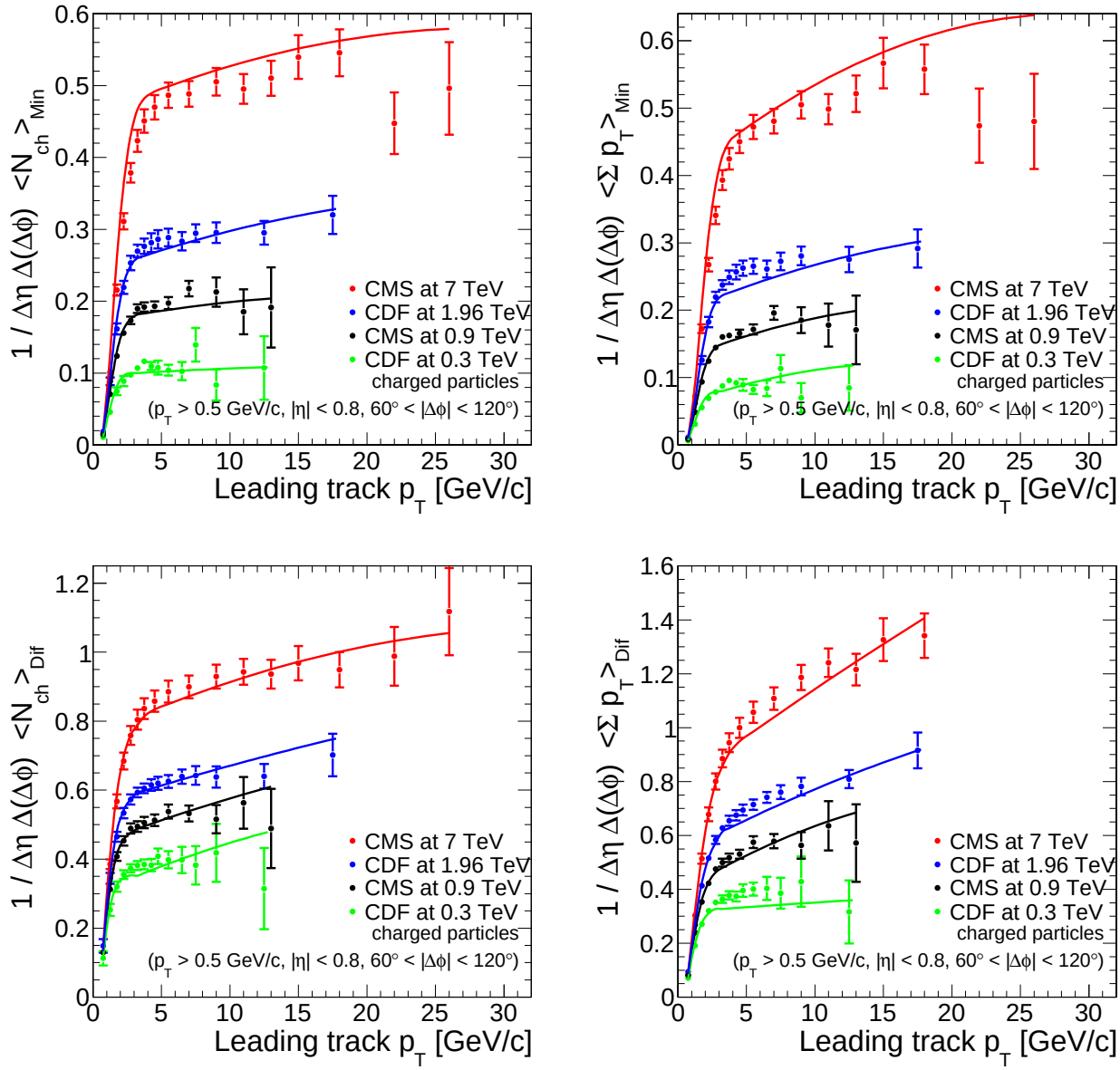

Figure 6-23. The hadronic activity at four center-of-mass energies in the min region and the difference in the activity between max and min.

6.8 Conclusions

- The hadronic activity associated with the UE has been studied at $7\ \mathrm{TeV}$ and $0.9\ \mathrm{TeV}$ at CMS. The two chosen observables, Nch and $\Sigma \mathrm{p_T}$, have been studied in various $\eta - \phi$ regions that emphasize different contributions to the UE dynamics.

- Two tunes that belong to the MC generator PYTHIA 6.4 are compared to data. Tune $\mathrm{Z1}$, the last tune produced in CMS using manual tuning, and tune $\mathrm{Z2^*}$ the first tune produced by CMS using PROFESSOR. There is a general success for both

tunes to simulate the data at the qualitative level. However, there is still room for improvement at the quantitative level, especially at the low values of leading track $\mathrm{p_T}$.

- Tune $\mathrm{Z2^*}$ gives close results to tune $\mathrm{Z1}$. It performs better for observable $\Sigma\mathrm{p_T}$ in most regions, particularly at the low leading track $\mathrm{p_T}$ values.

- The results obtained were compared with the results of other experiments at the same center-of-mass energy. Where the results seem to agree, with better statistics obtained by CMS that enables exploring regions with larger leading track $\mathrm{p_T}$.

- The results, alongside results from CDF, form basis to study the hadronic activity in various $\eta - \phi$ regions at four different center-of-mass energies. This forms basis to parametrize the energy dependence for each UE component more accurately for future tunes.

APPENDIX A
UNFOLDING

Unfolding is the procedure used to correct for any bias introduced to the observable due to experimental limitation (reconstruction, selection criteria, and physical limits etc.). The goal of unfolding is to restore the observable to truth or generator level spectrum that is measured with an ideal detector and an infinite event statistics. While unfolding is not essential for a discovery oriented analysis, it is important when the distribution itself is regarded as the set of parameters of interest, as is the case with the UE. In this analysis, we set the condition of having at least one charged track within the required cuts ($\mathrm{p_T} > 0.5\mathrm{GeV/c}$ and $|\eta| < 0.8$) to define an event at the generator level.

The method of choice is the bin-by-bin method. The main advantage is its simplicity and the transparency of the procedure while bin correlation is considered its main drawback. It is still used in many analyses in various collaborations. In this analysis we use the ratio of the observable (Nch and $\Sigma\mathrm{p_T}$ profile plots) at the generator over the detector level for a MC tune ($\mathrm{Z1}$) to correct the Data. Similar procedure was performed for MC tune $\mathrm{D6T}$ and the resulting distributions were close to the GEN level obtained directly from the MC.

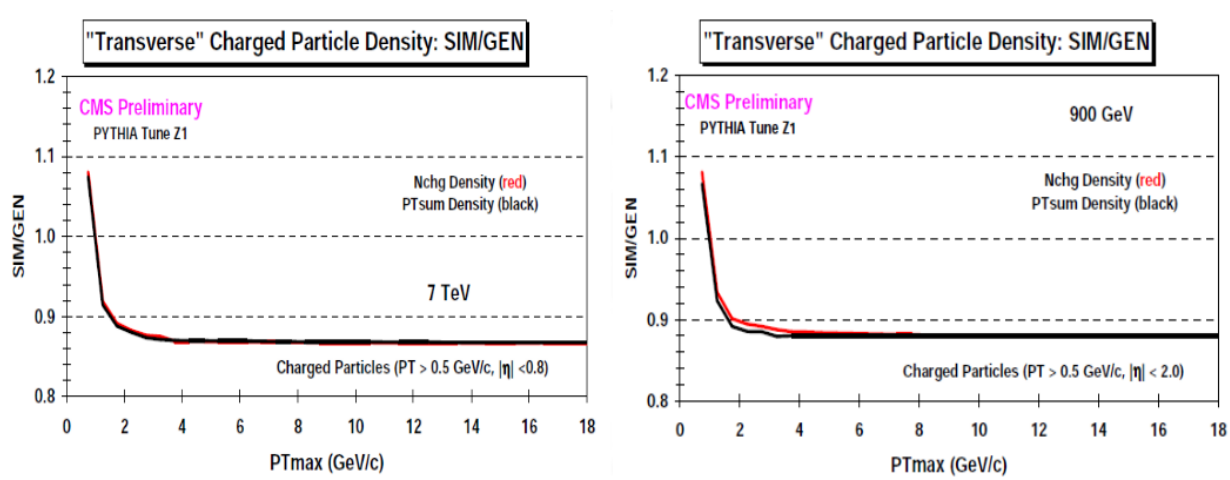

Figure A-1. The ratio of the MC of (PYTHIA 6, tune $\mathrm{Z1}$) simulation of the detector(SIM) level over the generator(GEN) level for the two observables: at $7\mathrm{TeV}$ (left panel) and at $0.9\mathrm{TeV}$ (right panel).

APPENDIX B
THE CROSS-SECTION FOR MPI

We start with Equation (2–1). From that we can write the differential cross-section as:

$$\frac{d\sigma(AB \to cd)}{dp_T^2} = \sum_{abcd} \int dx_a dx_b f_{a/A}(x_a, \mu_F^2) f_{b/B}(x_b, \mu_F^2) \frac{\hat{\sigma}_{abcd}(\alpha^2, \mu_R)}{dp_T^2} \ , \tag{B–1}$$

where

$$\frac{d\hat{\sigma}}{dp_T^2} = \frac{d\hat{\sigma}}{d\hat{t}} \frac{\pi\alpha_s^2}{\hat{s}^2} \frac{1}{\sqrt{1 - \frac{4p_T^2}{\hat{s}}}} \ , \tag{B–2}$$

at small scattering angles, t goes to zero, and the t-channel dominates. In this case the difference between quark and gluon interactions are due to their color factors:

$$\hat{\sigma}_{gg} : \hat{\sigma}_{qg} : \hat{\sigma}_{qq} = \frac{9}{4} : 1 : \frac{4}{9} \ , \tag{B–3}$$

The factors can be obtained from the gluon-gluon ($C_A = 3$) and gluon-quark ($T_F = \frac{1}{2}$), after including color-average factors: $\frac{1}{N^2-1} = \frac{1}{8}$ for the gluon and $\frac{1}{N} = \frac{1}{3}$ for the quark.

$$\sigma_{int} \approx \int_{p_{Tmin}}^{s/4} dp_T^2 \int_{4{p_T}^2/s}^{1} dx_a \int_{4{p_T}^2/sx_a}^{1} dx_b \frac{9}{2} \frac{\pi\alpha_s^2(p_T^2)}{\hat{s}^2\sqrt{1 - \frac{4p_T^2}{\hat{s}}}} F(x_A, p_T^2) F(x_B, p_T^2) f(\frac{\hat{u}}{\hat{t}}) \ , \tag{B–4}$$

$$\frac{d\hat{\sigma}}{dp_T^2} = \frac{8\pi\alpha_s^2(p_T^2)}{9p_T^4} \ , \tag{B–5}$$

$$\mathrm{F(x,Q^2) = g(x,Q^2) + \frac{9}{4}\Sigma_i[q_i(x,Q^2) + \bar{q}_i(x,Q^2)]} \ , \tag{B–6}$$

$$\sigma_{\mathrm{int}}(\mathrm{p_{Tmin}}) = \int\limits_{\mathrm{p_{Tmin}}}^{\mathrm{s/4}} \frac{\mathrm{d}\hat{\sigma}}{\mathrm{dp_T^2}} \mathrm{dp_T^2} \propto \frac{1}{\mathrm{p_{Tmin}^2}} \ , \tag{B–7}$$

APPENDIX C
SANITY CHECK PLOTS

Further sanity checks are shown in Figures C-1 and C-2. A good match between the MC and the data is observed in all the panels, which represent fundamental tracking variables.

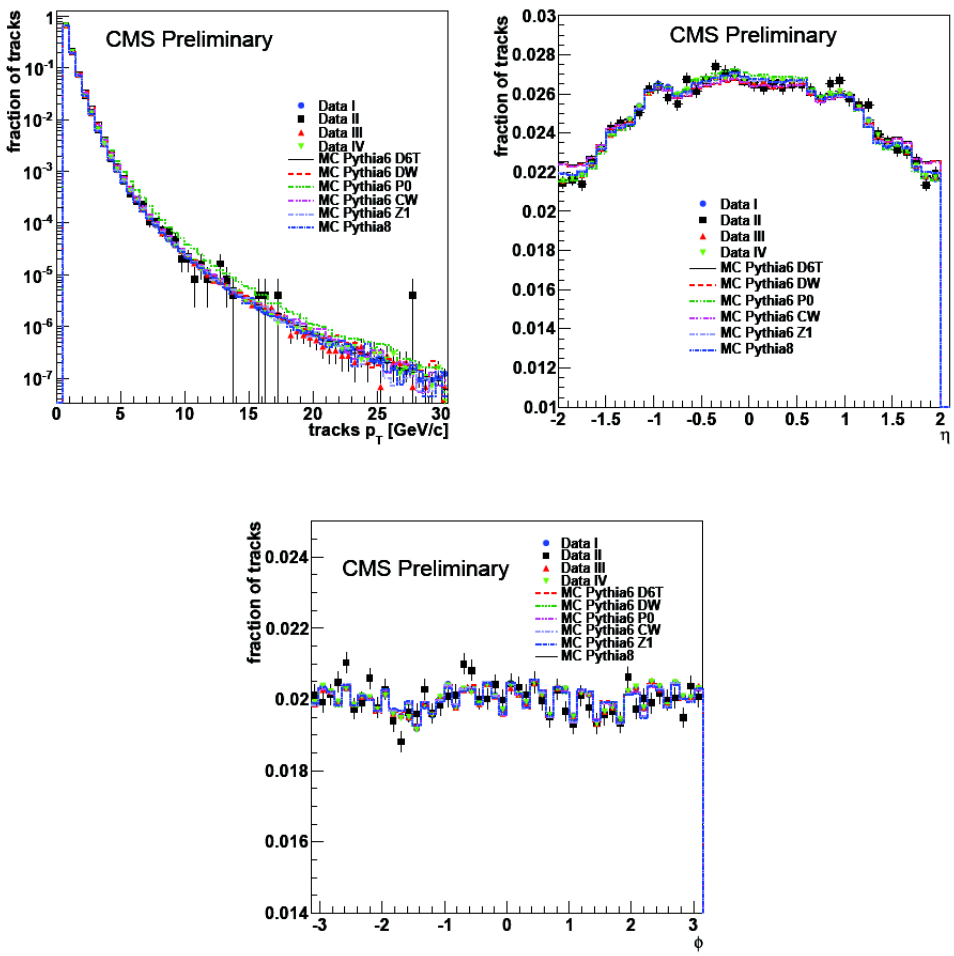

Figure C-1. $\mathrm{p_T}$, η and ϕ distributions for the reconstructed tracks at $7\ \mathrm{TeV}$. The unformity of the η and ϕ distributions and the success of the MC to simulate the data accurately indicate a good level of track reconstruction quality.

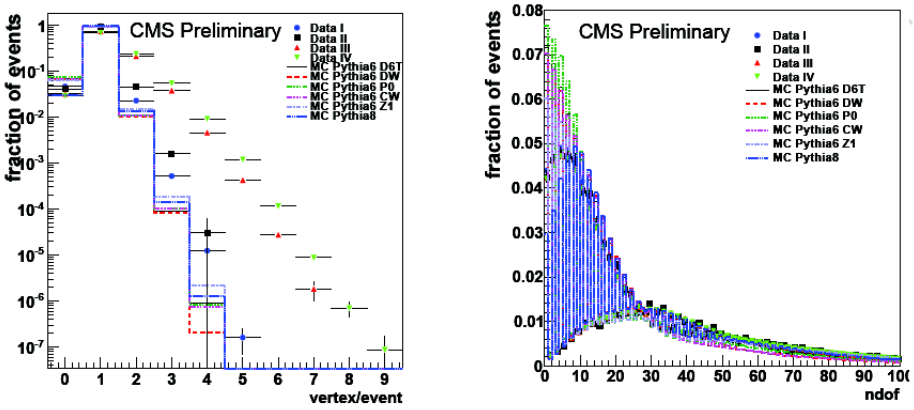

Figure C-2. Sanity check plots for event reconstruction: The left panel shows the number of offline vertices reconstructed for data and MC. First bin in this plot corresponds to the events in which no vertex is reconstructed. The right panel shows the ndof distribution of the real verteces.

APPENDIX D
SYSTEMATIC UNCERTAINTY

The simulation prediction is affected by inaccuracies in detector or beam condition modeling, so we estimate detector level systematic uncertainties by concentrating on the DATA/MC ratio of the observables. The systematic uncertainty associated is evaluated from the residuals with respect to a reference distribution: residuals $= (\mathrm{A} - \mathrm{B})/\mathrm{A}$, where A is the reference and B is the source. In addition to the direct evaluation it has also been considered an additional safety margin. The safety margin is introduced to take into account the bin by bin fluctuations, avoiding that the effect in bins with low statistics is underestimated by the fits.

The considered sources of detector level systematic uncertainties are the following:

- RECO to GEN correction.
- Pile-up contamination (comprising all the effects of high luminosity runs).
- Vertex selection.
- Tracking.
- Track selection and track cuts.

Table D-1 summarizes the systematic uncertainty values for the contributions mentioned above. The investigated sources of uncertainties are largely independent from each other, justifying the quadrature sum of the different contributions.

Table D-1. The main sources of systematic uncertainty.

Observable	tracking	track sel.	bg. cont.	vtx sel.	pile-up	$\sqrt{\mathrm{dz}^2 + \mathrm{d02}}$	MC model	total(%)
				$7\ \mathrm{TeV}$				
$< \mathrm{Nch} >$	1.02	1.4	0.8	0.8	0.7	2.8	2.0	3.6
$< \Sigma \mathrm{p_T} >$	1.04	1.1	0.8	1.2	0.7	2.3	2.0	3.7
				$0.9\ \mathrm{TeV}$				
$< \mathrm{Nch} >$	1.34	0.7	0.8	1.5	0	1.8	0.9	3.0
$< \Sigma \mathrm{p_T} >$	1.42	0.7	0.8	1.5	0	1.8	0.9	2.8

REFERENCES

[1] The CMS Collaboration, Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys.Lett. B **716**, 30 (2012). doi:10.1016/j.physletb.2012.08.021

[2] Bjorken, J. D. and Paschos, E. A., Inelastic Electron-Proton and γ-Proton Scattering and the Structure of the Nucleon. Phys. Rev. 185, 1975 (1969). doi:10.1103/PhysRev.185.1975

[3] G.P. Salam, Elements of QCD for hadron colliders. Lectures given at ESHEP (2009). CERN Yellow Report CERN-2010-002 arXiv:1011.5131v2

[4] http://www.isgtw.org/spotlight/go-particle-quest-first-cern-hackfest

[5] S. Bethke, The 2009 World Average of α_s. Eur. Phys. J. C64, 689 (2009).

[6] ALICE Collaboration, Measurement of inelastic, single- and double-diffraction cross sections in proton–proton collisions at the LHC with ALICE. Eur. Phys. J. C **73**, 2456 (2013). doi:10.1140/epjc/s10052-013-2456-0

[7] B. Combridge, J. Kripfganz, J. Ranft, Hadron Production At Large Transverse Momentum And QCD. Phys. Lett. B **70**, 234 (1977).

[8] J.M. Campbell, J.W. Huston, W.J. Stirling, Hard interaction of quarks and gluons: a primer for LHC physics. Rep. Prog. Phys., **70**, 89 (2007). doi:10.1088/0034-4885/70/1/R02

[9] P.Z. Skands, QCD for collider Physics. Lectures given at ESHEP (2010). CERN Yellow Report CERN-TH-2011-080. arXiv:1104.2863v2

[10] A.D. Martin, W.J. Stirling, R.S. Thorne, G. Watt, Parton distributions for the LHC Eur. Phys. J. C **63**, 189 (2009). doi:10.1140/epjc/s10052-009-1072-5

[11] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A.D. Poloa, ALPGEN, a generator for hard multiparton processes in hadronic collisions. J. High Energy Phys. **7**, 001, (2003). doi:10.1088/1126-6708/2003/07/001

[12] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, T. Stelzer, MadGraph 5 : Going Beyond. J. High Energy Phys. **06**, 128 (2011). doi:10.1007/JHEP06(2011)128

[13] V. V. Sudakov, Vertex parts at very high-energies in quantum electrodynamics. Sov. Phys. JETP. **3**, 65 (1956).

[14] T. Sjöstrand, S. Mrenna, P.Z. Skands, A Brief Introduction to PYTHIA 8.1. Comput.Phys.Commun. **178**, 852 (2008). doi:10.1016/j.cpc.2008.01.036

[15] T. Sjöstrand, S. Mrenna, P.Z. Skands, PYTHIA 6.4 physics and manual. J. High Energy Phys. **5**, 026 (2006). doi:10.1088/1126-6708/2006/05/026

[16] T. Gleisberg, S. Hoeche, F. Krauss, M. Schoenherr, S. Schumann, F. Siegert, J. Winter, Event generation with SHERPA 1.1. J. High Energy Phys. **02**, 007 (2009). doi:10.1088/1126-6708/2009/02/007

[17] M. Bähr et al., Herwig++ Physics and Manual. Eur. Phys. J. C **58**, 639 (2008). doi:10.1140/epjc/s10052-008-0798-9

[18] J. Beringer et al. (Particle Data Group), Phys. Rev. D **86**, 010001 (2012).

[19] T. Sjöstrand, Monte Carlo Generators for the LHC. Academic Training Lectures, CERN, Geneva (2005). http://agenda.cern.ch/fullAgenda.php?ida=a042790

[20] T. Sjöstrand, M. van Zijl, A multiple-interaction model for the event structure in hadron collisions. Phys. Rev. D **36**, 2019 (1987). doi:10.1103/PhysRevD.36.2019

[21] M. Bähr, J. M. Butterworth, M. H. Seymour, The Underlying Event and the Total Cross Section from Tevatron to the LHC. J. High Energy Phys. **01**, 065 (2009). doi:10.1088/1126-6708/2009/01/065

[22] L. Frankfurt, M. Strikman, C. Weiss, Transverse nucleon structure and diagnostics of hard parton-parton processes at LHC. Phys. Rev. D **83**, 054012 (2011). doi:10.1103/PhysRevD.83.054012

[23] R. Corke, T. Sjöstrand, J. High Energy Phys. **1**, 35 (2008). doi:10.1007/JHEP01(2010)035

[24] H. Schulz and P.Z. Skands, Energy Scaling of Minimum-Bias Tunes. Eur. Phys. J. C **71**, 1644 (2011). 10.1140/epjc/s10052-011-1644-z

[25] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, J. E. von Seggern, Systematic event generator tuning for the LHC. Eur. Phys. J. **C65**, 331 (2010). doi:10.1140/epjc/s10052-009-1196-7

[26] A. Buckley, J. Butterworth, D. Grellscheid, H. Hoeth, L. Lonnblad, J. Monk, H. Schulz, F. Siegert, Rivet user manual. MCnet/10/03 (2010). arXiv:1003.0694v8

[27] CMS Collaboration, Measurement of the Underlying Event Activity at the LHC with sqrt(s) = 7 TeV and Comparison with sqrt(s) = 0.9 TeV. J. High Energy Phys. **9**, 109 (2011). doi:10.1007/JHEP09(2011)109

[28] C. Lefevre, CERN FAQ - LHC the guide. CERN, Geneva (2013). CERN-Brochure-2008-001-Eng

[29] L. Evans, P. Bryant, LHC Machine. JINST **3**, S08001, (2008). doi:10.1088/1748-0221/3/08/S08001

[30] O.S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, P. Proudlock, LHC Design Report. 1. The LHC Main Ring. CERN-2004-003-V-1 (2008).

[31] E.A. Hillemanns et al., CMS General, CMS Detector posters, CMS Document 4227-v1 (2011).

[32] The CMS Collaboration et al, The CMS experiment at the CERN LHC. JINST **3**, S08004, (2008). doi:10.1088/1748-0221/3/08/S08004

[33] B.R. Martin, G. Shaw, Particle Physics. 104, John Wiley, United Kingdom (1997).

[34] CMS Collaboration, CMS Physics Technical Design Report. Volume I: Detector Performance and Software. CERN-LHCC-2006-001, Geneva (2006).

[35] A. Strandlie, Track reconstruction in the LHC experiments. 1st LHC Detector Alignment Workshop, CERN, Geneva, Switzerland, 89 (2006). doi:10.5170/CERN-2007-004

[36] R. Frühwirth, Application of Kalman filtering to track and vertex fitting. Nuc. Instrum. Methods Phys. Res., A **262**, 444 (1987). doi: 10.1016/0168-9002(87)90887-4

[37] R. Frühwirth, M. Regler, R.K. Bock, H. Grote, D. Notz, Data Analysis Techniques for High-Energy Physics. Cambridge Monographs On Particle Physics, Nuclear Physics and Cosmology **11**, 2nd edition, (2000). ISDN:0 521 63219 6

[38] P. Azzuri, B. Mangano, Optimal filtering of fake tracks. CMS Internal Note, CMS IN-2008/17 (2008).

[39] S. Cucciarelli, M. Konecki, D. Koltiński, T. Todorov, Track reconstruction, primary vertex finding and seed generation with the Pixel Detector. CMS note, 2006/26, (2006).

[40] T. Speer, K. Prokofiev, R. Früwirth, W. Waltenberger, P. Vanlaer, Vertex Fitting in the CMS Tracker. CMS Analysis Note, CMS AN-2006/032 (2006).

[41] D. Acosta et al., The underlying event in proton-proton collisions at 900 GeV CMS Analysis Note, CMS AN-2010/018 (2010).

[42] P. Bartalini et al., The underlying event in proton-proton collisions at 7 TeV CMS Analysis Note, CMS AN-2010/112 (2010).

[43] J. Incandela, S. A. Koay, R. Rossin, Primary Dataset definition procedures. CMS Internal Note, CMS IN-2009/02 (2008).

[44] T. Aumeyr, Evaluation of Oscilloscopes for the BPTX Read Out System - Tests in the ATLAS Trigger Electronics Lab. CMS Analysis Note, CMS AN-2007/049 (2007).

[45] L. Bagby, A. J. Bell, Y. J. Lee, G. Veres, The Beam Scintillation Counter Trigger System for CMS. CMS Detector Note, CMS DN-2010/018 (2010).

[46] T. Sjöstrand, P. Z. Skands, Multiple interactions and the structure of beam remnants. J. High Energy Phys. **03**, 053 (2004). doi:10.1088/1126-6708/2004/03/053

[47] ALICE Collaboration, Underlying Event measurements in pp collisions at sqrt(s) = 0.9 and 7 TeV with the ALICE experiment at the LHC. J. High Energy Phys. **7**, 116 (2012). doi:10.1007/JHEP07(2012)116

[48] CDF collaboration, Energy Dependence of Underlying Event (300 GeV, 900 GeV, and 1.96 GeV), http://www-cdf.fnal.gov/physics/new/qcd/QCD.html (access date 2013/11)

BIOGRAPHICAL SKETCH

Mohammed Khattab Zakaria was born in 1982 in Kuwait city, Kuwait. In 2000, he joined the University of Jordan majoring in physics. In 2004 he graduated with honors and joined the physics program at Creighton University. He received an award for outstanding scholarship upon the completion of his master's degree in 2007. Later in the same year, Mohammed joined the physics program at the University of Florida. Mohammed worked as a teaching assistant for the first 2 years at the physics department and he was awarded the Wayne Bomstad II award for being the distinguished teaching assistant of the year 2008. In 2009, Mohammed started working as a research assistant for his adviser Rick Field. Mohammed was involved in studying the underlying event activity associated with proton-proton collisions at the LHC. He was involved in the data analysis at CMS and the Monte Carlo modeling of the phenomena. During his years of graduate work, Mohammed was the recipient of the University Research Association fellowship that enable him to move to Fermilab to pursue his work. Mohammed was also the recipient of 2 LPC summer fellowships at the same facility. Mohammed received his Ph.D. in the fall of 2013.