UNDERLYING EVENTS AND SOFT QCD

PROBING PARTICLE PRODUCTION PHENOMENA IN HIGH ENERGY COLLISIONS AT THE LHC

WANG WEI YANG

(*B.Sc. (Hons.), NUS*)

DEPARTMENT OF PHYSICS

NATIONAL UNIVERSITY OF SINGAPORE

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN PHYSICS

Faculty of Science June 2016

I would like to dedicate this Thesis to my loving parents who have equipped me with all that I need in my life journey, and Hien who has supported and journeyed with me throughout the PhD candidature.

Declaration

I hereby declare that this thesis is my original work and it has been written by me in its entirety. I have duly acknowledged all the sources of information which have been used in the thesis.

This thesis has also not been submitted for any degree in any university previously.

WANG WEI YANG
(*B.Sc. (Hons.), NUS*)
June 2016

Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Chan Aik Hui Phil and Prof. Oh Choo Hiap for their patient guidance and support throughout my journey, providing me with wonderful learning opportunities in encouraging a critical and broad ranged thinking. Prof. Albert De Roeck for which this project would not have happened without him. I am very grateful to Dr. Xavier Janssen who was always willing to help, advise, and push me forward even when the going gets tough. Blessed to have colleagues at NUS HEP group: Mr. Ang Han Wei, Mr. Leong Qixiang, Mr. Shannon Seah, Mr. Jufri Setianegara, and many others with whom I had fun and enjoyable intellectual exchanges. I am thankful for the hospitality provided by the particle physics group at the University of Antwerp and for everyone who contributed to my enjoyable time there, including Prof. Pierre Van Mechelen, Prof. Nick Van Remortel, and Dr. Romain Rougny. Numerous colleagues and friends at CERN have given me the wonderful opportunity to work on an international stage with world class colleagues, and with whom many memories were made. Finally, the completion of the project is only possible through the support of the NUS Scholarship and Lee Foundation.

Table of Contents

Abstract **xi**

Preface **xiii**

List of Tables **xvii**

List of Figures **xix**

1 Quantum Chromodynamics **1**

1.1 Introduction 1

1.2 The Strong Force 3

1.2.1 QCD tests in e^+e^- and hadron colliders 6

1.3 Soft QCD 10

2 LHC and the CMS Collaboration **17**

2.1 Introduction 17

2.2 The LHC 18

2.3 The CMS Collaboration and Experiment 20

2.3.1 The CMS Detector 20

2.3.2 The CMS Trigger System 23

3 Selected Topics in Soft QCD **25**

3.1 Charged-particle Multiplicity Distribution 25

3.2 Angular Correlations 29

4 The Underlying Event **33**

4.1 Introduction 33

4.2 Methodology and Observables 35

4.3 Measurement of the UE at 2.76 TeV 39
4.3.1 Data and Monte Carlo samples 39
4.3.2 Event and track selection 47
4.3.3 Data Correction 49
4.3.4 Systematic uncertainties 62
4.3.5 Results 71
4.4 Measurement of the UE at 13 TeV 79
4.4.1 Results 80
4.5 Summary 83

5 Conclusion **89**

References **93**

Abstract

The theory of Quantum Chromodynamics (QCD) is well established and widely accepted to describe hadrons as composites of quarks strongly bound by gluon exchanges. The theory is described by an SU(3) Yang-Mills gauge theory and as such gluons are able to couple with each other, giving rise to profound consequences for the QCD coupling. The quarks and gluons are free and weakly coupled at small distances and the coupling strength increases with growing distance This gives rise to two salient features in QCD: Asymptotic freedom and confinement. The second feature is of great physical and mathematical significance but has not been shown theoretically even though it is consistent with experimental observations. It is a long standing problem in QCD where low momentum and large distance behaviour cannot be calculated perturbatively, even though all strongly interacting particles in experiments are confined and mostly produced through low momentum (soft) processes. Soft QCD phenomena is thus of huge phenomenological and experimental interest at the Large Hadron Collider (LHC) as they dominate most of the proton-proton interaction cross section and contribute to the background of other rare and interesting physics processes. The measurement of the underlying event forms an important component of soft QCD studies by quantifying some of the main features of soft QCD processes. This allows for improvements in the phenomenological modelling and description of soft QCD processes which in turn improves the precision measurement of standard model processes and searches of new physics. Recent highlights in various soft QCD phenomena, including multiplicity distributions, angular correlations, and the measurement of the underlying event at $\sqrt{s} = 2.76$ and 13 TeV with the CMS experiment at the LHC are presented. A steep rise in the underlying event activity is seen as a function of the jet transverse momentum up to about 8 GeV, followed by saturation. The underlying event activity also grows strongly with the centre-of-mass energy.

Preface

This thesis is written in fulfillment of the degree of Doctor of Philosophy in Physics with the National University of Singapore (NUS). The thesis' contents form part of my independent work spanning from experimental data analysis with the Compact Muon Solenoid (CMS) Collaboration at the Large Hadron Collider (LHC), to phenomenology which is traditionally strong in our group. A historic signing of the NUS-CMS (CERN) agreement in 2013 kicked off the collaboration between NUS and the CMS Collaboration, CERN. This has led to various publications, invited talks, and the pioneering work of this PhD project which serve to strengthen the capability of the HEP group with the various tools and techniques acquired, as well as to expand our collaborative efforts.

A year in total was spent with the particle physics group at the University of Antwerp and another year in CERN, where I did exploratory work on angular correlations and spearheaded the underlying event analysis at 2.76 TeV, and another underlying event analysis at 13 TeV using data collected with an upgraded CMS detector in a different run condition. The analyses involve scope and strategy definition, data cleaning, object selection, Monte Carlo event production, data correction, estimation of systematic uncertainties, and writing of various papers. The rest of the PhD candidature was spent in NUS with a focus on the phenomenological description of multiplicity distributions.

This thesis is guided by the overarching theme of soft Quantum Chromodynamics (QCD) which encompass the above mentioned topics. The contents are organised as such: Chapter 1 introduces the standard model and high energy collider physics, emphasising the behaviour of QCD and soft QCD. Chapter 2 describes the LHC and the CMS detector. Chapter 3 covers the topic of multiplicity distributions and angular correlations. Chapter 4 discusses the underlying event analyses, the result of the very first NUS-CMS collaboration, followed by a conclusion in chapter 5.

List of Publications:

1. W. Y. Wang, S. Seah, J. Setianegara, A. H. Chan, and C. H. Oh, "A weighted GMD model for multiplicity distributions at LHC energies," *(preprint arXiv)*, 2016

2. W. Y. Wang, "Recent measurement of underlying events," *Proceedings of the 7th International workshop on Multiple Partonic Interactions at the LHC*, p. 39, 2016

3. W. Y. Wang, S. Seah, J. Setianegara, A. H. Chan, and C. H. Oh, "The weighted GMD model for multiplicity distributions at LHC energies," *Proceedings of the XLV international symposium on multiparticle dynamics*, *(accepted)*, 2016

4. W. Y. Wang, "Underlying event and correlation results from CMS," *Proceedings of the XLV international symposium on multiparticle dynamics*, *(accepted)*, 2016

5. CMS Collaboration, "Underlying event measurements with leading particles and jets in pp collisions at $\sqrt{s} = 13$ TeV," *CMS Physics Analysis Summary*, vol. CMS-PAS-FSQ-15-007, 2015

6. S. Bansal, A. H. Chan, D. Ciangottini, L. Fano, R. Field, C. H. Oh, D. Rank, and W. Y. Wang, "Underlying event via leading track and track jet at 13 TeV," *CMS Analysis Note*, vol. CMS-AN-2015/184, 2015

7. CMS Collaboration, "Measurement of the underlying event activity using charged-particle jets in proton-proton collisions at 2.76 TeV," *JHEP*, vol. 09, p. 137, 2015, *[arXiv:1507.07229]*

8. W. Y. Wang, A. H. Chan, and C. H. Oh, "Measurement of the underlying event activity using charged-particle jets in proton-proton collisions at $\sqrt{s} = 2.76$ TeV," *Proceedings of the international conference on massive neutrinos*, p. 283, 2015

9. CMS Collaboration, "Measurement of the underlying event activity using charged particle jets in proton-proton collisions at $\sqrt{s} = 2.76$ TeV," *CMS Physics Analysis Summary*, vol. CMS-PAS-FSQ-12-025, 2014

10. S. Bansal, A. H. Chan, X. Janssen, C. H. Oh, N. Remortel, and W. Y. Wang, "Measurement of the Underlying Event using track-jets in *p*-*p* Collisions at 2.76 TeV," *CMS Analysis Note*, vol. CMS-AN-2012/282, 2014

11. W. Y. Wang, "Measurement of the underlying event activity using charged-particle jets in proton-proton collisions at 2.76 TeV," *Proceedings of the 6th international workshop on multiple partonic interactions at the LHC*, *[arXiv: 1506.05829]*, p. 21, 2014

12. W. Y. Wang, Q. Leong, W. K. Ng, A. Dewanto, A. H. Chan, and C. H. Oh, "Deformed Coherent States for Multiparticle Production Mechanism," *Proceedings for the Conference in Honour of the 90th Birthday of Freeman Dyson*, p. 400, 2014

Invited Speaker:

1. 23rd Annual Astronomy Retreat, Kuantan, Malaysia, 2016
2. Collider, ArtScience Museum, Singapore, 2016
3. Conference on New Physics at the Large Hadron Collider, Singapore (NTU), Singapore, 2016
4. 7th International workshop on multiple partonic interactions at the LHC (MPI@LHC), Trieste (ICTP), Italy, 2015
5. XLV international symposium on multiparticle dynamics, Wildbad Kreuth, Germany, 2015
6. International conference on massive neutrinos, Singapore (NTU), Singapore, 2015
7. Institute of Physics Singapore (IPS) Annual March Meeting, Singapore (NTU), Singapore, 2015
8. 6th international workshop on multi partonic interactions (MPI@LHC), Krakow (PAU), Poland, 2014
9. Conference in Honour of the 90th Birthday of Freeman Dyson, Singapore (NTU), Singapore, 2013
10. 1st IAS-CERN Workshop on Particle Physics and Cosmology, Singapore (NTU), Singapore, 2013
11. International Conference on Flavour Physics in the LHC Era, Singapore (NTU), Singapore, 2010

List of Tables

4.1 Summary of the parameters of the Monte Carlo generator tunes. 46
4.2 Summary of event and track selection at detector level. 48
4.3 Phase space definition for charged jet and charged particles at generator level. 56
4.4 Summary of the systematic uncertainties (in percentage) due to various sources and in the various transverse regions. 70

List of Figures

1.1 "Standard model of Elementary Particles" by MissMJ - PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group. 3

1.2 The energy evolution of electron-positron (filled circles) and hadron (filled squares) colliders [1]. The energy of the hadron colliders are scaled down by a factor of 6 to 10 to represent the energy of the constituent quarks and gluons. 4

1.3 The three primary and complementary (anti) colours: red (anti-red/cyan), green (anti-green/magenta), and blue (anti-blue/yellow) [2, 3]. The various modes of interactions between the different colour charges are illustrated as field lines connecting the charges. Their combinations describe the baryons (three colours) and mesons (two colours). 5

1.4 Illustration of high energy e^+e^- annihilation into (left) many-particle final-state that appears to be clustered as two "jets" of particles. The production rates, energy, and angular distributions of the (right) single quarks or gluons that underlie each cluster of particles can be calculated in QCD. 7

1.5 Observation of (left) di- and (right) tri-jet events by the CMS detector in pp collisions at the LHC. 8

1.6 Feynman diagram for the process (left) $e^+e^- \to \mu^+\mu^-$ and (right) $e^+e^- \to q\overline{q}$ interacting through a virtual photon γ. 9

1.7 World data on the ratio R of the total cross-section $\sigma(e^+e^- \to hadrons)$ to $\sigma(e^+e^- \to \mu^+\mu^-)$, as a function of the collision centre-of-mass energy $\sqrt{s}$. The data points are compared to the naive quark-parton model prediction (dashed curve), and 3-loop pQCD calculation (solid curve) [4, 5]. 10

1.8 Measurements of the coupling constant α_s as a function of the respective energy scale of Q [6]. Open symbols indicate (resummed) NLO and filled symbols NNLO QCD calculations used in the respective analysis. The curves are the QCD predictions for the combined world average value of $\alpha_s(M_Z)$ with a 4-loop approximation. 11

1.9 The 4 stages of a $e^+e^- \to hadrons$ process: (1) electroweak radiation and annihilation, (2) development of the quark-antiquark system with gluon radiation described by pQCD, (3) fragmentation/hadronisation process where the non-perturbative regime sets in and the conversion of quarks/gluons into hadrons are described phenomenologically, (4) decay of unstable hadrons. . 12

1.10 Next to leading order PDFs (MSTW 2008) [7], showing the distribution of partons with longitudinal momentum fraction x_L probed at energy scales of (left) $Q^2 = 10$ and (right) $Q^2 = 10^4$ GeV^2. The different types of partons u, d, s, c, b, g, and their corresponding antiparticles are labelled. The coloured bands show the 1-sigma uncertainty. Note that the gluon distribution is scaled by $1/10$. 13

1.11 The scattering cross-section of several important processes compared to the total pp cross-section as a function of the centre-of-mass collision energy [8]. The vertical axis on the right shows the equivalent event rate per luminosity of $L = 10^{33}$ cm^{-2}s^{-1}. . 14

2.1 The CERN complex of accelerators and experiments. Past accelerators like the LINAC, BOOSTER, PS, and SPS are integrated with the 27 km LHC ring serving as intermediate accelerators. . 19

2.2 (Top left) A picture of the CMS detector in an opened up state showing its cross section. (Top right) Schematic drawing of the CMS detector, identifying the main sub-detectors and components that makes it up. (Bottom) A segment of the CMS detector cross-section showing the layers of sub-detectors and types of particles that it detects. 21

2.3 A slice of the CMS detector showing the coordinate system convention. The x, y, and z-axes are shown together with some typical η values. 22

3.1 (Left) Charged-particle multiplicity distribution for pp collisions at $\sqrt{s} = 0.9$ TeV across all rapidity windows as measured by the CMS Collaboration [9]. The results are compared to the same measurements done by the UA5 Collaboration [10] for $p\overline{p}$ collisions and by the ALICE Collaboration [11] for pp collisions. (Right) Comparison of the charged-particle multiplicity distributions for $p\overline{p}$ collisions at $\sqrt{s} = 200$, 540, 900 GeV, and a phenomenological description by the NBD [12]. Data points are measured by the UA5 Collaboration [13, 10]. 26

3.2 Comparison of various phenomenological models (left) GMD and WGMD, and (right) PYTHIA8 MC with and without MPI, against MD data at $\sqrt{s} = 7$ TeV. 28

3.3 The 2D two-particle angular correlation functions for (top left) 2.76 TeV *PbPb* and (top right) 5.02 TeV *pPb* collisions using charged-particle pairs with $1 < p_T < 3$ GeV in events where the particle multiplicity is within $220 \leq N_{trk}^{offline} < 260$. The same distribution is shown for pp collisions at 13 TeV with (bottom left) $N_{trk}^{offline} < 35$ and (bottom right) $N_{trk}^{offline} \geq 105$. The right and bottom figures have sharp peaks that are truncated to emphasise the other correlation structures. 30

3.4 The integrated associated yield as a function of the multiplicity $N_{trk}^{offline}$ for the near-side of the correlation function for pp collisions at $\sqrt{s} = 7$ and 13 TeV, pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and $PbPb$ collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Error bars denote statistical uncertainties while the shaded areas and boxes represent the systematic uncertainties. 32

4.1 Illustration of correlations in azimuthal angle $\delta\phi$ relative to (*l*eft) the direction of the leading track, or to (right) the leading track-jet. 36

4.2 Average scalar sum p_T density of charged particles as a function the azimuthal angle difference $\Delta\phi$ between each particle relative to the leading particle with (left) $p_T > 1$ GeV and (right) $p_T > 2$ GeV. The data points (filled circles) are compared to several PYTHIA MC tunes (lines) which went through full detector simulation. 37

4.3 Illustration of the separation of the overall transverse region into the transMAX and transMIN regions based on the activity in the region. 38

4.4 Venn diagram showing the set of minimum bias and Jet20 triggered events ordered in p_T^{jet}. The overlapping region is the sample of events where both triggers fired. The p_T^{jet} for which the co-triggered event is 100% that of the minimum bias event sample within the p_T^{jet} slice is set as the threshold. 41

4.5 (Top left) Leading jet p_T, (top right) efficiency $\left(PRSC\_Jet20 \times \frac{N_{MinBias+Jet20}}{N_{MinBias}}\right)$ of Jet20 trigger w.r.t. minimum bias trigger, (bottom left) ratio $\left(\frac{PRSC\_Jet20 \times N_{Jet20}}{PRSC\_MB \times N_{MinBias}}\right)$ of all Jet20 triggered events w.r.t minimum bias triggered events and (bottom right) efficiency $\left(PRSC\_Jet40 \times \frac{N_{Jet20+Jet40}}{N_{Jet20}}\right)$ of Jet40 trigger w.r.t. Jet20 trigger. In the top left plot, events are weighted according to trigger prescales. 42

4.6 Comparison of track properties for data and Monte Carlo prediction PYTHIA6 Z2 and PYTHIA8 4C (top left) p_T, (top right) η, (centre left) ϕ, (centre right) d_0/σ_{d_0}, (bottom left) d_z/σ_{d_z}, (bottom right) σ_{p_T}/p_T. Bottom panel in each plot shows the ratio of the MC prediction and the measurement to quantify the level of agreement. 50

4.7 Efficiency (left column) and fake rate (right column) for reconstructed tracks as a function of track (top) p_T, (centre) η, and (bottom) ϕ obtained using PYTHIA6 MC. 51
4.8 Correlation between leading jet reconstructed with the SisCone and Anti-k_T algorithms (top) jet p_T, (centre) jet η and (bottom) jet ϕ using PYTHIA6 MC. 52
4.9 Comparison of UE activity as a function of leading track-jet clustered using SisCone and Anti-k_T algorithms for (left) transverse particle density and (right) transverse Σp_T density . Bottom panel in both plots shows the ratio UE activity as a function of leading track-jet clustered using SisCone and Anti-k_T algorithms. 53
4.10 Comparison of track-jet properties for data and prediction PYTHIA6 Z2 (left) p_T, and (right) η. Bottom panel in each plot shows the ratio of the MC prediction and the measurement to quantify the level of agreement. 53
4.11 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) total transverse particle density and (right) total transverse Σp_T density as a function leading track-jet p_T. 54
4.12 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transMAX particle density and (right) transMAX Σp_T density as a function leading track-jet p_T. 54
4.13 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transMIN particle density and (right) transMIN Σp_T density as a function leading track-jet p_T. 55
4.14 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transDIF particle density and (right) transDIF Σp_T density as a function leading track-jet p_T. 55
4.15 Correlation matrices for jet properties at reconstructed and generator level; (top) leading jet p_T, (centre) leading jet η and (bottom) leading jet ϕ. These are obtained using PYTHIA6 MC sample. 57
4.16 Correlation matrices for jet properties at reconstructed and generator level; (left) leading jet η and (right) leading jet ϕ with the z-axis in logarithmic scale. These are obtained using PYTHIA6 MC sample. 58

4.17 Correlation matrices for track properties at reconstructed and generator level; (left) total transverse track multiplicity and (right) total transverse Σp_T. These are obtained using PYTHIA6 MC sample. 58

4.18 Correlation matrices for track properties at reconstructed and generator level; (left) transMAX track multiplicity and (right) transMAX Σp_T. These are obtained using PYTHIA6 MC sample. 58

4.19 Correlation matrices for track properties at reconstructed and generator level; (left) transMIN track multiplicity and (right) transMIN Σp_T. These are obtained using PYTHIA6 MC sample. 59

4.20 Correlation matrices for track properties at reconstructed and generator level; (left) transDIF track multiplicity and (right) transDIF Σp_T. These are obtained using PYTHIA6 MC sample. 59

4.21 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transverse particle density and (right) transverse Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample. 59

4.22 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transMAX particle density and (right) transMAX Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample. 60

4.23 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transMIN particle density and (right) transMIN Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample. 60

4.24 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transDIF particle density and (right) transDIF Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample. 61

4.25 χ^2 of unfolding at different iterations for PYTHIA6 MC unfolded with an independent sample of PYTHIA6 MC; comparison is done for the various transverse regions of (left) particle density and (right) Σp_T density. 61

4.26 Comparison of unfolded results (left) particle density, (right) Σp_T density obtained by Bayesian method and simple bin-by-bin correction. Bottom plot shows the ratio of profile distributions corrected using Bayesian method and bin-by-bin correction. . . 62

4.27 Unfolded PYTHIA8 4C profile distributions of (left) transverse particle density and (right) transverse Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.012 with $\chi^2/NDF = 14.26/10$ and (right) 1.016 with $\chi^2/NDF = 7.93/10$. 63

4.28 Unfolded PYTHIA8 4C profile distributions of (left) transMAX particle density and (right) transMAX Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.008 with $\chi^2/NDF = 9.58/10$ and (right) 1.009 with $\chi^2/NDF = 4.81/10$. 64

4.29 Unfolded PYTHIA8 4C profile distributions of (left) transMIN particle density and (right) transMIN Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.011 with $\chi^2/NDF = 11.41/9$ and (right) 1.014 with $\chi^2/NDF = 5.96/9$. 64

4.30 Unfolded PYTHIA8 4C profile distributions of (left) transDIF particle density and (right) transDIF Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 0.979 with $\chi^2/NDF = 5.11/10$ and (right) 0.980 with $\chi^2/NDF = 7.27/10$. 65

4.31 Unfolded data profile distributions of (left) transverse particle density and (right) transverse Σp_T density where Jet20 events are weighted by 1.05 and compared to the normal distribution. Bottom panels represent the ratio of the biased over un–biased distributions. 66

4.32 Unfolded data profile distributions of (left) transverse particle density and (right) transverse Σp_T density where Jet20 events are weighted by 0.95 and compared to the normal distribution. Bottom panels represent the ratio of the biased over un–biased distributions. 66

4.33 Estimation of the pileup effect with (left) the distribution of number of vertices with a Poisson fit and (right) the Δz distribution of 2 vertices. 68

4.34 Summary of the systematic uncertainties for the profile distributions panel (left) for transverse particle and (right) for transverse $\sum p_T$ densities. 71

4.35 Summary of the systematic uncertainties for the profile distributions panel (left) for transMAX particle and (right) for transMAX $\sum p_T$ densities. 71

4.36 Summary of the systematic uncertainties for the profile distributions panel (left) for transMIN particle and (right) for transMIN $\sum p_T$ densities. 72

4.37 Summary of the systematic uncertainties for the profile distributions panel (left) for transDIF particle and (right) for transDIF $\sum p_T$ densities. 72

4.38 Measured (top) particle density, and (bottom) Σp_T density, in the transverse region relative to the leading charged-particle jet in the event ($|\eta| < 2$, $60° < |\Delta\phi| < 120°$), as a function p_T^{jet}. The data (symbols) are compared to various MC simulations (curves). The ratios of MC simulations to the measurements are shown in the bottom panels. The inner error bars correspond to the statistical uncertainties, and the outer error bars represent the statistical and systematic uncertainties added in quadrature. . . . 74

4.39 Measured (top) particle density, and (bottom) Σp_T density, in the transMAX region ($60° < |\Delta\phi| < 120°$, relative to the leading charged-particle jet in the event, with maximum UE activity), as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38. 75

4.40 Measured (top) particle density, and (bottom) Σp_T density, in the transMIN region ($60° < |\Delta\phi| < 120°$, relative to the leading charged-particle jet in the event, with minimum UE activity), as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38. 76

4.41 Measured transDIF activity (see text for its definition) for (top) particle density, and (bottom) Σp_T density, as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38. 77

4.42 Comparison of UE activity at $\sqrt{s} = 0.9$, 2.76, and 7 TeVfor (top) particle density, and (bottom) Σp_T density, as a function of p_T^{jet} [14, 15]. The data (symbols) are compared to various MC simulations (curves). The definition of the error bars is the same as for Fig. 4.38. 78

4.43 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average particle densities with the various simulations as a function of p_T^{jet}. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band. . 81

4.44 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average energy densities with the various simulations as a function of p_T^{jet}. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band. . 82

4.45 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average particle densities with the various simulations as a function of p_T. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band. . 84

4.46 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average energy densities with the various simulations as a function of p_T. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band. . 85

4.47 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN particle densities with various simulations at $\sqrt{s} = 0.9$, 2.76, 7, and 13 TeVfor the top left plot, and $\sqrt{s} = 2.76$ and 13 TeVfor the rest as a function of p_T^{jet}. 86

4.48 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN energy densities with various simulations at $\sqrt{s} = 0.9$, 2.76, 7, and 13 TeVfor the top left plot, and $\sqrt{s} = 2.76$ and 13 TeVfor the rest as a function of p_T^{jet}. 87

Chapter 1

Quantum Chromodynamics

1.1 Introduction

On 10 September 2008, the Large Hadron Collider (LHC) started up for the first time, becoming the world's largest and most powerful particle accelerator with a design specification able to reach a maximum collision energy of 14 TeV. The physics program at CERN, the European Organisation for Nuclear Research, aims to answer some of the most fundamental questions about nature with the LHC. What is the universe made of and how did it begin?

In order to answer these questions, physicists at the LHC pump protons or lead ions into accelerators that speed them up to 99.9999991% of c (the speed of light) and collide them at very high energies in an attempt to create conditions similar to the beginning of the universe. Under such conditions, a multitude of new and exotic sub-atomic particles are produced. As a consequence of the mass-energy relation $E = mc^2$, we know that the more energetic a collision the more mass can be produced from the collision. This not only translates to the production of more particles, but also to the creation of exotic new particle species with higher masses. All these particles being created at the LHC would have come into their brief existence at some point during the infancy of our universe. By studying these particles, high energy physicists are in effect peering into the photo album of her early developmental stage to discover answers as to what makes up our universe, how she was born and how she grew to be who she is today.

Of all the particle species produced at colliders, only the proton, neutron, and electron seems to factor into our daily lives, making up the atoms and the elements that form our basis of existence. However remote these exotic particles may seem, their existence hints at the underlying laws in which they interact; their creation during the Big Bang suggest that their participation in the early evolution of the universe were vital in determining the current state of our universe and existence.

Despite this, their obvious lack of participation in our lives and the multitude which were being discovered in the early days of particle physics posed serious challenges to a systematic way in which they could be organised and having a theory that could describe their behaviour. Such were the problems that early particle physicists faced.

Today, the standard model is able to accurately and precisely describe all the known sub-atomic particles and how they interact. Of the four fundamental forces through which the particles interact, three of which are described by the standard model. The electromagnetic, strong, and weak interactions (forces) are all incorporated in the standard model as paradigms of Quantum Field Theory, a framework that describes particles as excited states of an underlying physical field.

According to the standard model, our universe is made up of only a dozen fundamental particles, six quarks and six leptons. The interacting forces between these particles are themselves described by three types of mediating particles: the photon, the gluon, and the W and Z bosons. The theory also also includes the Higgs boson, the only particle of its kind that gives mass to the other particles. These are summarised in figure 1.1.

The discovery of previously unknown particle species or new species that have been predicted to exist predicates the validation of theories and search for new phenomena. Thus the construction of bigger and higher power colliders often lends itself as a straightforward way to produce these particles for research. This has led to generations of colliders over the years, moving towards higher energies. Figure 1.2 shows the evolution of collider energies across the years, up to LHC energies at the terascale. The most recent discovery of the Higgs boson (2013) at the LHC was a smashing success for high energy physics and the standard model, confirming the existence of the last missing particle required by the model.

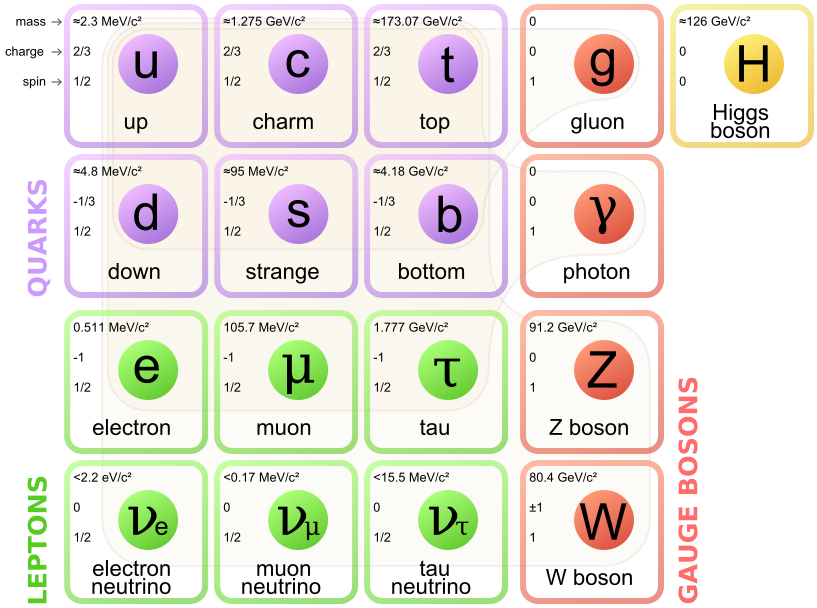

Figure 1.1 "Standard model of Elementary Particles" by MissMJ - PBS NOVA, Fermilab, Office of Science, United States Department of Energy, Particle Data Group.

1.2 The Strong Force

In the world that we are used to, the electromagnetic force manifests itself in almost all aspects of our day-to-day living from electronics to the chemistry of the elements. These are described by the theory of Quantum Electrodynamics (QED). We know from our daily experience and high school physics that the strength of the electric and magnetic force decreases with increasing distance between charges or magnets, and increases as they are brought closer to each other. In the terminology of collider physics, the interaction strength increases with collision energy, which corresponds to a large momentum transfer and small separation. As the collision energy increases, the coupling strength can theoretically run to infinity (Landau pole) at a finite energy scale of about 10^{274} TeV. Running into the Landau pole may signify that QED is incomplete and would have to be extended or superseded by a new theory. Such theories are of phenomenological (modelling and description of phenomena not describable from first principles) interest but not currently testable nor pressing since even the LHC's maximum design energy only reaches about 14 TeV. At such energy scales, QED is well tested and is able to make very accurate predictions.

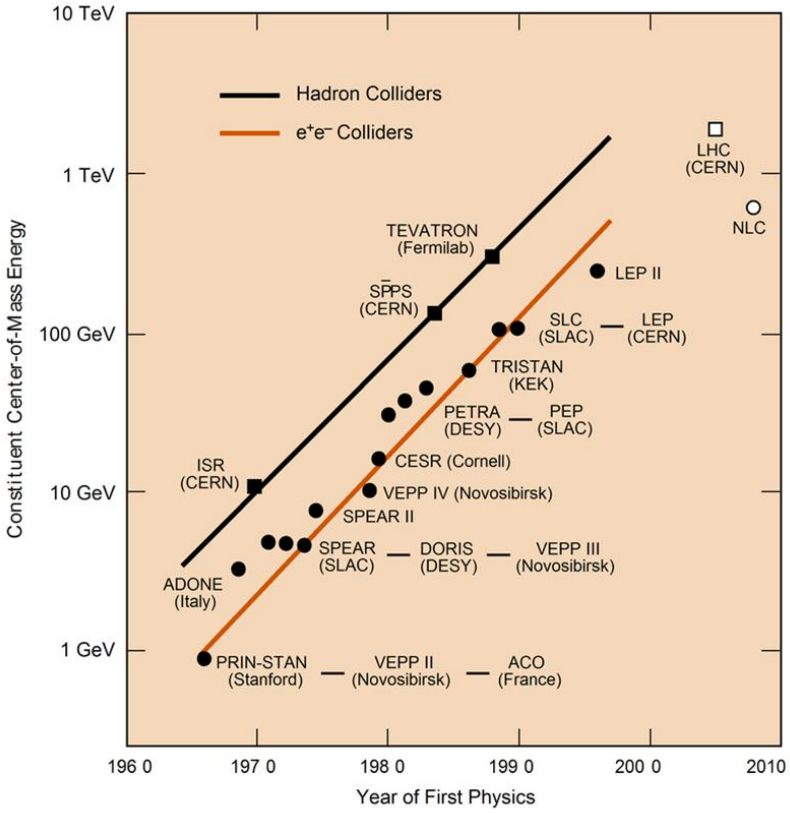

Figure 1.2 The energy evolution of electron-positron (filled circles) and hadron (filled squares) colliders [1]. The energy of the hadron colliders are scaled down by a factor of 6 to 10 to represent the energy of the constituent quarks and gluons.

In the standard model, the quarks and leptons are ascribed with electric charges and interact electromagnetically. While the charged leptons hold a unit of electronic charge, the quarks have fractional electric charges of either $\pm 1/3$ or $\pm 2/3$. Interestingly, the quarks hold another kind of charge called the colour charge. This means that the quarks also interact with each other through the strong force, which is described by the theory of Quantum Chromodynamics (QCD). In analogy to the three primary colours, there are three colour charges (and three anti-charges): red (anti-red), green (anti-green), and blue (anti-blue). The interactions between the colour charges are then mediated by gluons. Figure 1.3 summarises the types of interactions between the colour charges.

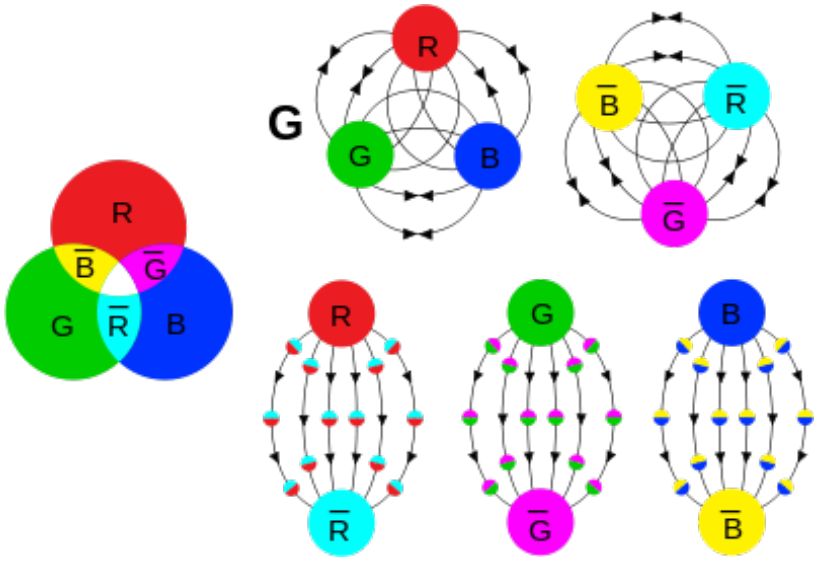

Figure 1.3 The three primary and complementary (anti) colours: red (anti-red/cyan), green (anti-green/magenta), and blue (anti-blue/yellow) [2, 3]. The various modes of interactions between the different colour charges are illustrated as field lines connecting the charges. Their combinations describe the baryons (three colours) and mesons (two colours).

The theory of QCD is described as a Yang-Mills gauge theory, with eight gauge fields predicted by the non-abelian gauge group SU(3). This means that the gauge fields which we call gluons are able to interact with each other. As a consequence, QCD has starkly different properties from what we are used to in QED. The theory has two main properties:

1. Asymptotic freedom: When the separation between the quarks and gluons are small, they interact very weakly. As the separation decreases, the

quarks and gluons behave more like free particles and form a state called quark gluon plasma (QGP).

2. Confinement: The strong force between quarks tends to increase with separation. This suggests that quarks tend to be bound together. In trying to separate them, the gluon field energy gets large enough to create new quark-antiquark pairs that in turn tend to be bound. As a result, free quarks can never be observed.

The second property is of great physical and mathematical significance but has not been verified theoretically, even though it is consistent with experiments. No free quark has ever been observed; only hadrons, the colour-singlet state (i.e. combination of 3 quarks with different colours or 2 quarks with a colour and anti-colour) of quark combinations.

Nevertheless, perturbative QCD (pQCD) is able to make definitive predictions about how quarks and gluons can be observed in an indirect manner. In high energy collisions, the quarks and gluons behave as free particles that interact at short distances and scatter energetically off each other to form cascades of well collimated gluons and quark-antiquark pairs through the process of bremsstrahlung radiation. This is well described by calculations in pQCD. However as the quarks and gluons lose their energy through the cascade, the shower quickly enters a non-perturbative regime at an energy scale of $\Lambda_{QCD} \approx 0.2$ GeV. At this stage, the quarks and gluons transform into equally well collimated "jets" of hadrons in a process called hadronisation which is not describable through pQCD. Thus at large distances (> 1 fm), the quarks and gluons are confined as jets of hadrons which are observable. The properties of the underlying parent quarks and gluons, such as spin, flavour, and colour can then be inferred.

1.2.1 QCD tests in e^+e^- and hadron colliders

Particle colliders are an essential tool for high energy particle physics, and the credibility of QCD was established through the successful operation of various high energy colliders. As seen in 1.2, particle colliders can be classified into two types: leptonic colliders that smash electrons and positrons or hadronic colliders that collide protons, antiprotons, and heavier ions. Electron-positron colliders offer the benefit of having "cleaner" final-states (or resulting constituents) since

they do not have substructure (like hadrons do) and only interact through the electromagnetic and weak forces. Hadronic colliders on the other hand sheds light on the substructure of the hadrons and the strong force.

Existence of jets

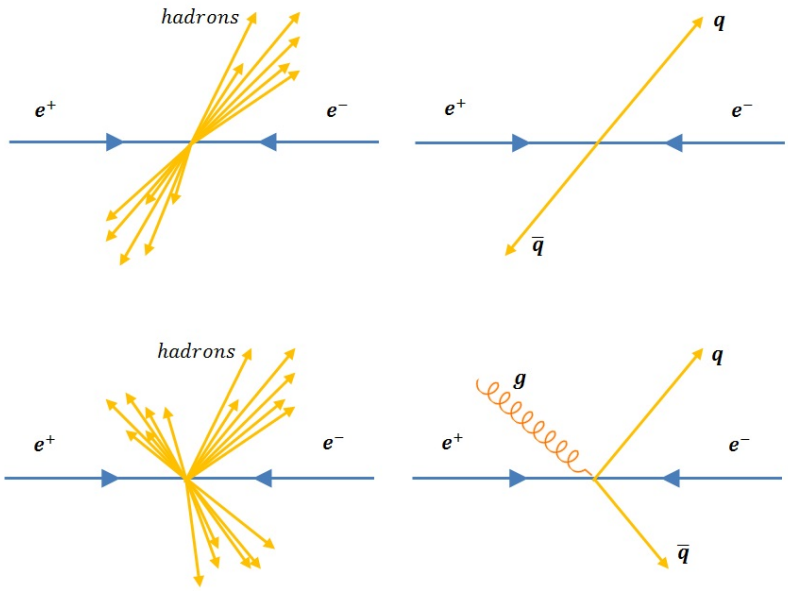

Figure 1.4 Illustration of high energy e^+e^- annihilation into (left) many-particle final-state that appears to be clustered as two "jets" of particles. The production rates, energy, and angular distributions of the (right) single quarks or gluons that underlie each cluster of particles can be calculated in QCD.

Results from early QCD tests in e^+e^- colliders provided crucial evidence in support of QCD. One such example was the observation of jets in the annihilation of e^+e^- into a quark (q) and antiquark ($\overline{q}$) pair in the process (Fig. 1.4 top):

$$e^+e^- \rightarrow q\overline{q} \rightarrow 2\,jets. \tag{1.1}$$

This process was predicted in the early 1970s [16–18]. Shortly after, the first di-jet events were observed by the SPEAR collider [19, 20] at the Stanford Linear Accelerator Center (SLAC). This was followed by the first evidence of gluons at PETRA in the Deutsches Electronen-Synchrotron (DESY) [21–24] in the form of three-jet events where a gluon (g) is radiated from one of the quarks (Fig. 1.4

bottom):

$$e^+e^- \rightarrow q\bar{q}g \rightarrow 3\,jets. \tag{1.2}$$

Again, this was predicted theoretically [25] a few years before the observation of gluon jets. The process through which the gluon is radiated, results in a characteristic pattern in the observed final-state similar to that observed in photon bremsstrahlung ($e^+e^- \rightarrow e^+e^-\gamma$) in QED. Through such a comparison, it became clear that the gluon is generated through a bremsstrahlung process. Today, events containing two and three jets (Fig. 1.5) are frequently observed by the Compact Muon Solenoid (CMS) detector with proton-proton (pp) collisions at the LHC.

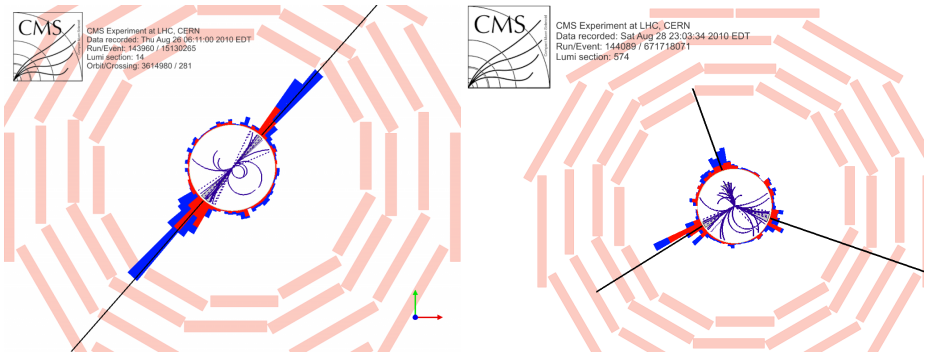

Figure 1.5 Observation of (left) di- and (right) tri-jet events by the CMS detector in pp collisions at the LHC.

The Drell-ratio R

The di-jet production process (Eq. (1.1)) can be represented in terms of a Feynman diagram as shown in figure 1.6 (right), where q and $\overline{q}$ represents the quark and antiquark flavours which can be u, d, s, c, b, or t and their corresponding antiquarks. In this picture, the e^+e^- annihilates to form a virtual photon γ which then creates a quasi-free quark-antiquark pair. Since the quarks carry colour charge, they are prevented from being completely free. In a simplistic jet production model, the quark-antiquark pair directly hadronises to form a jet pair.

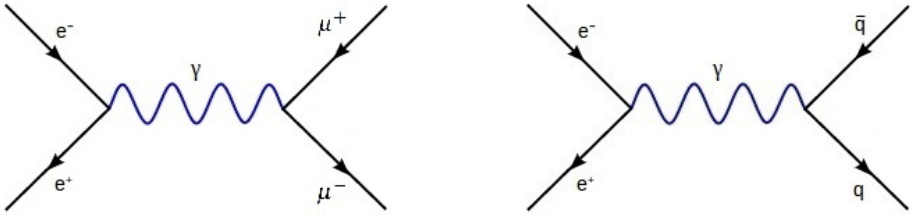

Figure 1.6 Feynman diagram for the process (left) $e^+e^- \to \mu^+\mu^-$ and (right) $e^+e^- \to q\bar{q}$ interacting through a virtual photon γ.

The simple quark model [16–18] is able to make definitive calculations of the Drell-ratio R supported by experimental measurements (Fig. 1.7). In the model,

$$R \equiv \frac{\sigma\left(e^+e^- \to hadrons\right)}{\sigma\left(e^+e^- \to \mu^+\mu^-\right)} = N_C \sum_f Q_f^2, \tag{1.3}$$

where $\sigma\left(e^+e^- \to hadrons\right)$ is the total cross-section (total production probability, roughly speaking) of the annihilation to form hadrons and $\sigma\left(e^+e^- \to \mu^+\mu^-\right)$ is the total cross-section for $\mu^+\mu^-$ production. The right hand side of equation (1.3) can be calculated for number of colours $N_C = 3$ and quark charge Q_f for flavours $f = u, d, s, c, b$. This gives a value of 2 (for $f = u, d, s$), increasing to $10/3$ (for $f = u, d, s, c$) above the $e^+e^- \to c\bar{c}$ (J/ψ) threshold and $11/3$ (for $f = u, d, s, c, b$) beyond the $e^+e^- \to b\bar{b}$ (Υ) threshold. Clearly, the experimental data supports $N_C = 3$. For any other value of N_C, the calculated ratio will not match experimental measurements.

Measurement of the running QCD coupling constant, α_s

The measurement of the running coupling constant α_s provides another strong test of QCD. As the theory describes a force that is asymptotically free at small distances and confinement at large distances, this translates to a coupling strength that is small (free) in an energetic collision and grows larger (confined) as the energy decreases. This means that the coupling constant runs as a function of the momentum Q, transferred between the probe and the target particle.

Figure 1.8 shows various measurements of the coupling constant at different Q. At large Q, $\alpha_s(Q)$ gets smaller and measurements are more precise, as expected from asymptotic freedom. As Q tends towards 1 GeV, $\alpha_s(Q)$ grows very quickly and the level of precision drops. The measurements compare well

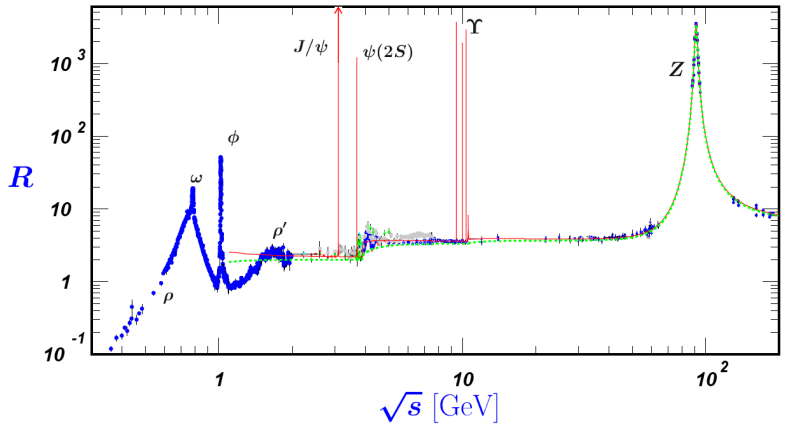

Figure 1.7 World data on the ratio R of the total cross-section $\sigma(e^+e^- \to hadrons)$ to $\sigma(e^+e^- \to \mu^+\mu^-)$, as a function of the collision centre-of-mass energy $\sqrt{s}$. The data points are compared to the naive quark-parton model prediction (dashed curve), and 3-loop pQCD calculation (solid curve) [4, 5].

with a 4-loop pQCD approximation extrapolated down to $Q = 1$ GeV where $\alpha_s(Q) = 0.5$. The measurements provide strong evidence for asymptotic freedom and pQCD predictions. For $Q < 1$ GeV, perturabative calculations start to become unfeasible, signalling the setting in of confinement.

1.3 Soft QCD

Soft QCD is the non-perturbative regime of QCD, where the energy scale of the interaction is very small. In collider physics the momentum Q serves as a proxy for the energy scale of the interaction in which the strength of the strong coupling constant α_s runs, as seen in figure 1.8. However in hadronic collisions, the transverse momentum of jets (p_T^{jet}) or particles (p_T) serve as more useful proxies for the interaction energy scale. Thus soft QCD concerns the study of low p_T/soft particles or jets.

To illustrate the transition between the perturbative and non-perturbative regimes of QCD physics, the simplest description is provided by that of the process $e^+e^- \to hadrons$ as shown in figure 1.9. In this description, stage 1 of the process involves only electroweak processes ending with the production of a virtual photon γ or Z boson. In stage 2, the γ/Z produces a high p_T quark-antiquark pair which radiates additional gluons in processes that are calculable by pQCD.

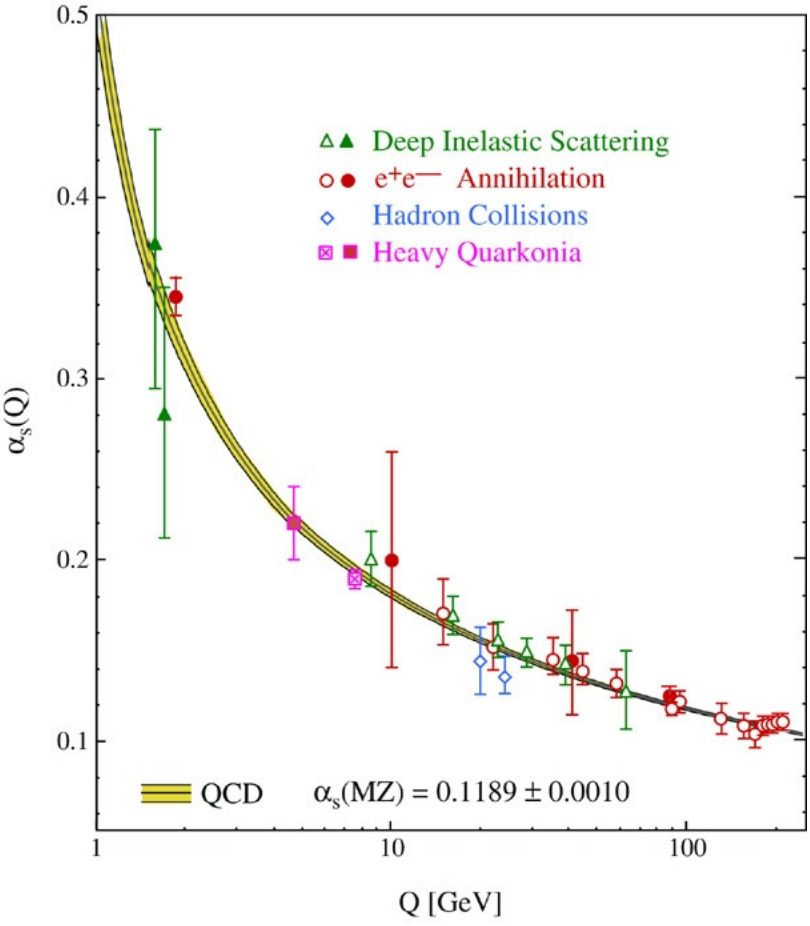

Figure 1.8 Measurements of the coupling constant α_s as a function of the respective energy scale of Q [6]. Open symbols indicate (resummed) NLO and filled symbols NNLO QCD calculations used in the respective analysis. The curves are the QCD predictions for the combined world average value of $\alpha_s(M_Z)$ with a 4-loop approximation.

As the quark radiates gluons, its p_T drops as it evolves down the energy scale. When $p_T \sim \Lambda_{QCD}$, the development enters the non-perturbative stage 3, where hadronisation (or fragmentation) occurs. In this stage, various phenomenological models of hadronisation are required to describe the conversion from the quarks and gluons into hadrons. A basic hypothesis is Local Parton Hadron Duality (LPHD), where the local momentum and quantum numbers are conserved when the partons (quarks and gluons) turns into hadrons. After hadronisation, the resulting hadrons that are unstable undergoes decay in stage 4.

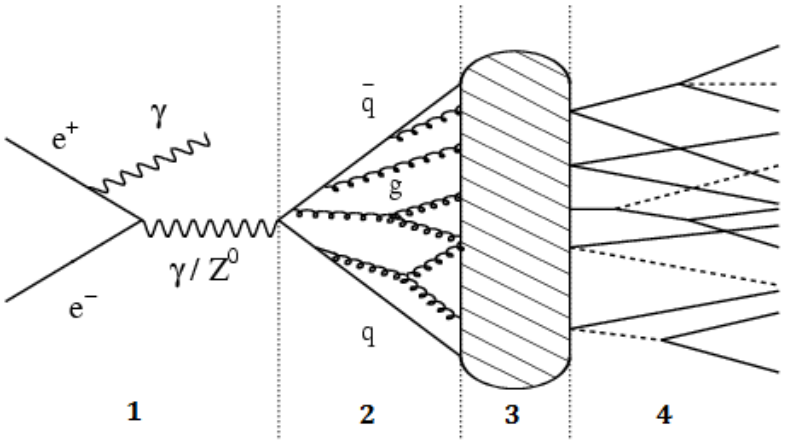

Figure 1.9 The 4 stages of a $e^+e^- \rightarrow hadrons$ process: (1) electroweak radiation and annihilation, (2) development of the quark-antiquark system with gluon radiation described by pQCD, (3) fragmentation/hadronisation process where the non-perturbative regime sets in and the conversion of quarks/gluons into hadrons are described phenomenologically, (4) decay of unstable hadrons.

In hadronic collisions like pp collisions at the LHC, the situation is complicated by the fact that hadrons are made up of partons bound in a confined state. Deep inelastic scattering experiments have revealed the structure and behaviour of the partonic constituents in measurements of Parton Distribution Functions (PDF). An example is shown in figure 1.10, where the number density of the various types of partons holding longitudinal momentum fraction x_L of the proton are shown. The valence quarks' (u and d) contribution to the total momentum of the proton is clearly large. However, a huge number of gluons holding a small x_L also contributes significantly to the overall proton momentum. Typically, the gluons can hold up to 50% of the total proton momentum. As the energy scale at which the PDF is probed is increased from $Q^2 = 10$ to 10^4 GeV^2, the number of sea quarks and gluons increase by so much at low x_L that the valence quarks are

overshadowed. This is understood as the contribution from vacuum polarisation being resolved by the more energetic probe. Thus the PDFs have to be taken into account in very high energy pp collisions since a larger fraction of partons carry small x_L, leading to collisions that tend to be softer.

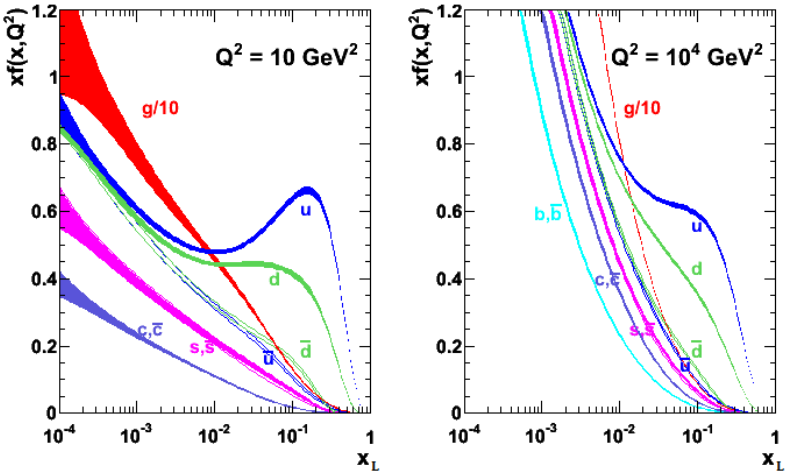

Figure 1.10 Next to leading order PDFs (MSTW 2008) [7], showing the distribution of partons with longitudinal momentum fraction x_L probed at energy scales of (left) $Q^2 = 10$ and (right) $Q^2 = 10^4$ GeV^2. The different types of partons u, d, s, c, b, g, and their corresponding antiparticles are labelled. The coloured bands show the 1-sigma uncertainty. Note that the gluon distribution is scaled by $1/10$.

Soft QCD processes contribute significantly to the total cross-section of pp collisions. This is clearly seen in figure 1.11, where the cross-section of several important standard model processes are shown alongside the total pp cross-section, σ_{tot}. Processes that are abundant at LHC energies like bottom quark production are a 100 times smaller than σ_{tot}, while rare processes like Higgs production is expected to happen in at most 1 in 10 billion collisions at a centre-of-mass energy of $\sqrt{s} = 7$ TeV. The large difference between σ_{tot} and the other processes is attributed to soft QCD [8]. As a result soft QCD processes factor into most measurements as QCD background, affecting the level of precision obtainable.

A good understanding of soft QCD is thus important for precise measurements of standard model processes as well as searches for new physics. In any given proton beam bunch crossing at the LHC, multiple pp collisions can take place simultaneously (pileup, PU). Since each PU collision is independent and dominated by soft QCD, these processes form an irreducible background to other

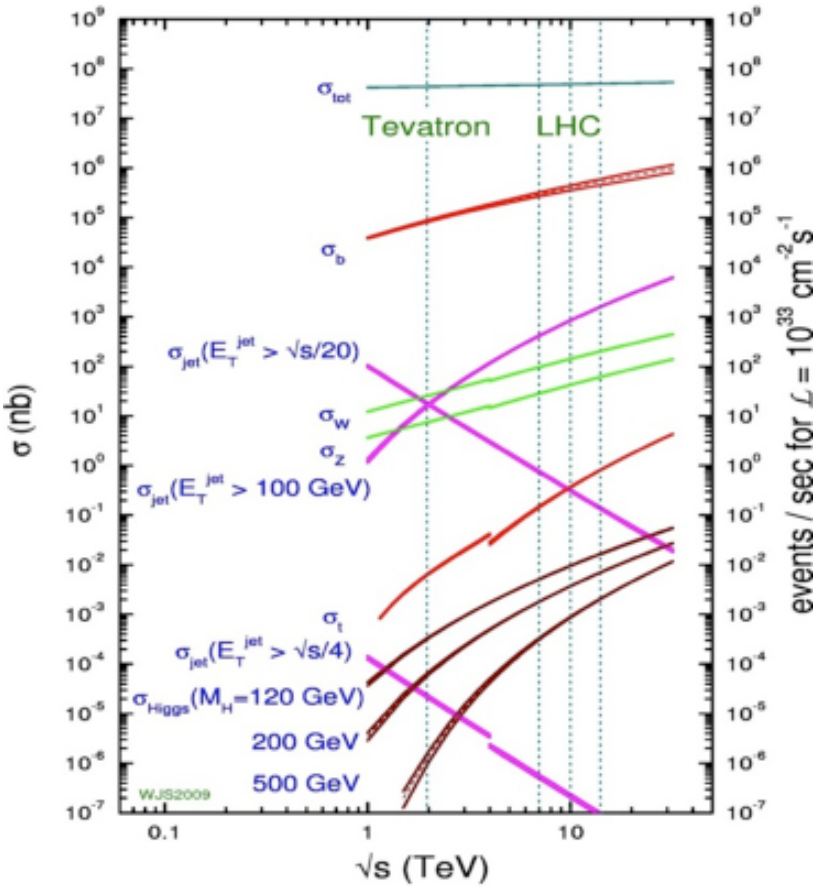

Figure 1.11 The scattering cross-section of several important processes compared to the total pp cross-section as a function of the centre-of-mass collision energy [8]. The vertical axis on the right shows the equivalent event rate per luminosity of $L = 10^{33}$ cm^{-2}s^{-1}.

processes with similar signature as well as affecting measurements of isolated high p_T leptons or photons. These measurements in turn form the basis for searches of interesting/rare standard model or beyond standard model processes like Higgs or Supersymmetric (SUSY) particle production. Thus the PU collisions magnify the effects of soft QCD, affecting the precision level obtainable for such measurements.

Since soft QCD processes cannot be adequately calculated by pQCD methods, other ways must be devised in order to improve our understanding. In general two phenomenological approaches have proved to be very useful in providing insights to soft QCD physics. The first is a top down approach of looking at global characteristics of soft radiation through statistical modelling of the final-state distributions, before proceeding to interpret the parameters of the model. Since particle production in high energy collisions have an inherent quantum statistical nature and soft QCD radiation is most easily produced in abundance due to its large cross-section, this approach provides a natural way of describing the data. The second method is a bottom up approach of using Monte Carlo (MC) event generators. Due to the exponential increase of computational power over the years through various advancements including the usage of grid computing, this approach has become very popular especially within the experimental particle physics community. These event generators simulate individual particle collisions and the evolution stages as illustrated by figure 1.9. By generating billions of such individual collision events, various results from experimental observations of real collisions can be reproduced by processing the MC generated events in exactly the same way. These MC generators are programmed to generate standard model processes using perturbative methods, including pQCD. In order to simulate soft QCD processes, various phenomenological models can be built in to complement pQCD.

With the above strategies in mind, the following chapters will present a non-exhaustive list of studies which aim to improve our understanding of soft QCD physics. An experimental and phenomenological discussion of the following topics will be given:

- Charged-particle multiplicity distributions
- Angular correlations
- Underlying events (UE).

These topics are selected based on their interrelationship and a personal contribution towards the discussed measurements and analyses, as part of this PhD research. Special focus is placed on the measurement of the UE which forms a major component of my research contribution done with measurements performed at the CMS detector at the LHC.

Chapter 2

LHC and the CMS Collaboration

2.1 Introduction

The Conseil Européen pour la Recherche Nucléaire (CERN) or European Organisation for Nuclear Research, is the research organisation that houses the largest particle physics research laboratory in the world. Situated in the laboratory is the world's most powerful particle collider, the Large Hadron Collider (LHC).

The CERN was established on 29th September 1954, born out of a world that was war devastated, with countries forming international organisations like the UN and UNESCO to strengthen relations between people. Several prominent European scientists including Niels Bohr and Louis de Broglie were among the pioneers who put forth the proposals leading to the CERN convention that was eventually ratified by 12 countries.

Since its formation, CERN has housed numerous particle accelerators from the SynchroCyclotron (SC) in 1957, which was no more than 20 m in circumference, to the LHC in 2008, with a ring that is 27 km in circumference. Several important achievements were made at the colliders. These include the first creation and isolation of antihydrogen, discovery of neutral currents, the W and Z bosons, pentaquarks, and the 125 GeV particle consistent with the Higgs boson. Apart from physics, CERN also contributed to other fields like detector technologies, medical physics, and the creation of the World Wide Web.

The LHC began operation on 10th September 2008, but was forced to stop after the system failed due to faulty electrical connection to a magnet, causing mechanical damage and release of liquid Helium into the tunnel. Repairs took

14 months and the first run eventually started again on 20th November 2009, setting a new world record with collision centre-of-mass energies of 2.36 TeV. Currently, the LHC runs at centre-of-mass energies of 13 TeV.

2.2 The LHC

The CERN complex of accelerators is shown in figure 2.1. The LHC itself is housed in a 27 km tunnel 100 m underground, across the Swiss-Franco border near Geneva city. As a pp collider, hydrogen atoms are stripped of their electrons before being pumped into the LINAC accelerator for the initial kick. The protons are then passed to the BOOSTER and then the PS where they are separated into bunches. The proton bunches are then further accelerated in the SPS where they are finally injected into the LHC.

The LHC itself comprises of two concentric beampipes maintained in ultra-high vacuum for the circulation of the proton bunches. The Niobium Titanium (NbTi) dipole magnets that are held at 1.9 K provide a magnetic field of 8.3 T (or 100000 times stronger than the Earth's magnetic field), allowing the bunches of protons to circulate in opposite directions. To create collisions, the proton bunches are made to collide at 4 interaction points (IP) where the 2 beampipes intersect. During heavy ion runs, lead ions are also passed through the same process.

The LHC was designed to have pp collisions reach a maximum centre-of-mass energy of 14 TeV. Thus far, the LHC has operated at 0.9, 2.36, 2.76, 5, 7, 8, and 13 TeV. During each run, around 2808 bunches are made to circulate in each ring, with a 25 ns bunch spacing, and 115 billion protons per bunch. With these parameters, the instantaneous luminosity can be calculated with

$$L \equiv \frac{1}{\sigma}\frac{dN}{dt} = f n_b \frac{N_b^2}{A}, \tag{2.1}$$

where f is the revolution frequency of the proton bunches around the LHC rings, n_b is the number of bunches per beam, N_b is the number of protons per bunch, and A is the overlap between the crossing bunches at the IP. Throughout the LHC operation, the luminosity has been adjusted and ramped up at various times to approach the design goal of $L = 10^{34}$ cm^{-2}s^{-1}. Experimentally, a higher luminosity means a higher chance of pp interaction per bunch crossing, boosting

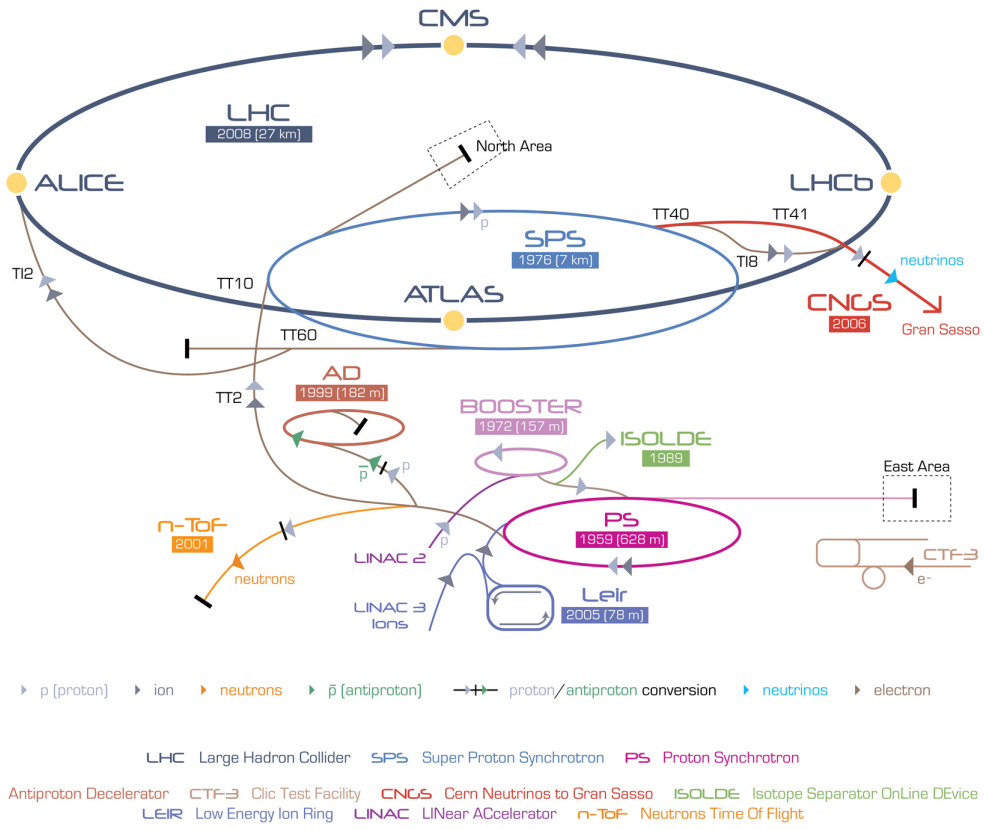

Figure 2.1 The CERN complex of accelerators and experiments. Past accelerators like the LINAC, BOOSTER, PS, and SPS are integrated with the 27 km LHC ring serving as intermediate accelerators.

the amount of data that can be collected. This is especially important in searches for rare processes like Higgs production. Such high luminosities lead to more than 25 independent pp interactions in the same bunch crossing (pileup, PU).

There are 4 large experiments running at the LHC with their respective detectors installed where the 4 IP are. These are the ALICE, ATLAS, CMS, and LHCb. As the name ALICE (A Large Ion Collider Experiment) suggests, the experiment focuses on the study of heavy ion collisions. In such collisions, the density is expected to be high enough to produce the quark gluon plasma, a state of matter that was created in the early universe, lasting a few millionths of a second. The physics programme of the LHCb (LHC beauty) collaboration focuses on studying how matter came to dominate antimatter as the universe evolved, through looking at interactions involving the beauty/bottom quark. The ATLAS (A Toroidal LHC Apparatus) and the CMS (Compact Muon Solenoid) are large general purpose colliders with wide ranging physics interests, studying various standard model processes and searches for new physics.

2.3 The CMS Collaboration and Experiment

The CMS collaboration consists of an international collaboration spanning over 42 countries and 182 institutes. There are about 4300 people working in this collaboration, including about 2680 who are physicists and students. In 2013, the National University of Singapore (NUS) and CMS collaboration signed the NUS-CMS (CERN) Expression of Interest. Since then, NUS has officially participated in CMS activities under the scheme.

2.3.1 The CMS Detector

The CMS collaboration runs the CMS detector and analyses the data collected from it. The detector (Fig. 2.2 (top left)) weighs 12500 tonnes, is 21.6 m long, and 14.6 m tall. It comprises of several cylindrical layers (Fig. 2.2 (top right)) of sub-detector components that act like filters designed to stop, track, or measure the different types of particles produced by the collision. Figure 2.2 (bottom) shows a segment of the cross-section, revealing the various sub-detector components: silicon tracker, electromagnetic calorimeter (ECAL), hadronic calorimeter (HCAL), superconducting solenoid, iron yoke, and the muon chambers. The

whole detector is permeated by a magnetic field of up to 3.8 T. The iron yoke which weighs 12000 tonnes, serves to contain and concentrate the magnetic field for high precision momentum measurements.

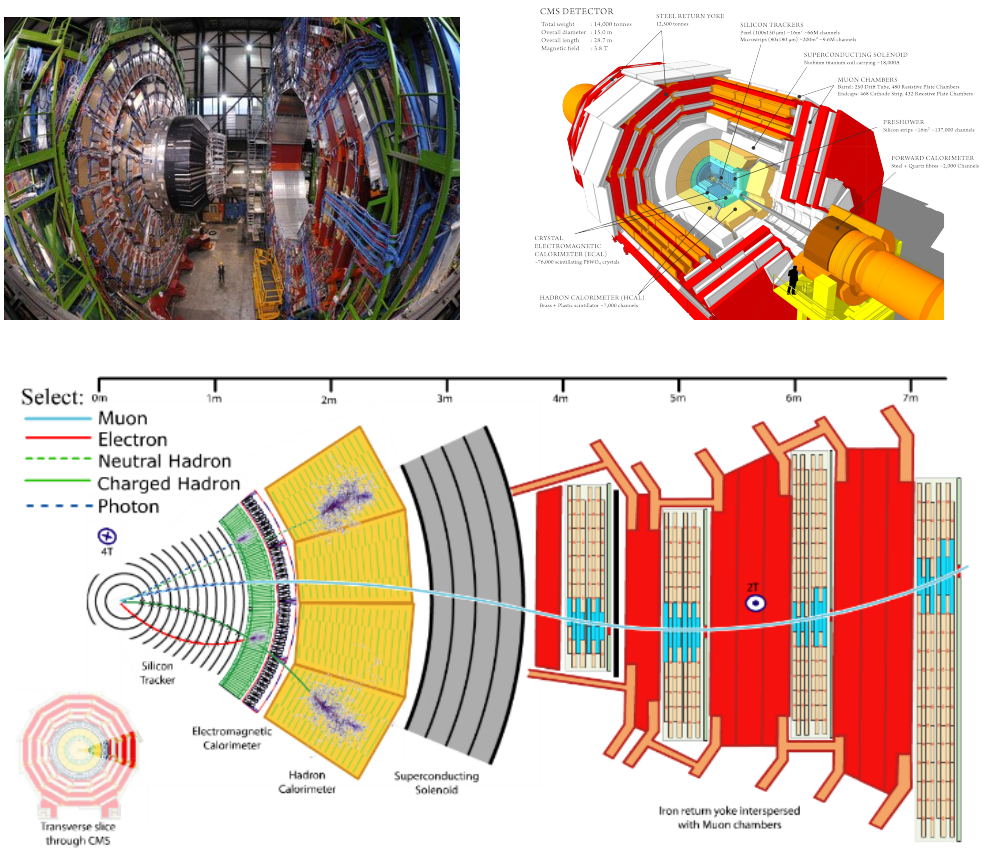

Figure 2.2 (Top left) A picture of the CMS detector in an opened up state showing its cross section. (Top right) Schematic drawing of the CMS detector, identifying the main sub-detectors and components that makes it up. (Bottom) A segment of the CMS detector cross-section showing the layers of sub-detectors and types of particles that it detects.

The first layer in which particles emerging from the collisions encounter is the silicon pixel and strip tracker. As charged-particles move through the tracker, their position is picked up and the paths that they take can be reconstructed. Such reconstructed paths are called tracks. As the tracks are curved due to the magnetic field, the momentum of the particles are also revealed.

The next two layers are the ECAL and HCAL. These calorimeters measure the energy of the particles by stopping them. The ECAL is designed to measure the energy of photons and electrons precisely while the HCAL will stop hadrons.

Muons will pass through these layers and their paths are then picked up by the muon chambers, which is the last layer. Neutrinos, being weakly interacting and neutral, pass through the detector without leaving any trace.

The CMS detector uses a coordinate system with the IP as the origin of the axis, as illustrated in figure 2.3. The x-axis is pointed towards the radial direction of the cylindrical detector, the y-axis directed upwards, and the z-axis in the direction of the beam running clockwise within the LHC ring. The angle θ is defined with respect to the z-axis in the yz-plane, and ϕ with respect to the y-axis in the xy-plane. The pseudorapidity variable

$$\eta = \frac{1}{2}\ln\left(\frac{|p|+p_L}{|p|-p_L}\right) = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{2.2}$$

is more commonly used in place of θ. In the limit where the measured outgoing particles are moving close to the speed of light, $|p| \approx E$ and therefore $\eta \approx y$, where y (not the same as the y-axis) is the experimental definition of rapidity

$$y = \frac{1}{2}\ln\left(\frac{E+p_L}{E-p_L}\right). \tag{2.3}$$

The experimental rapidity y differs slightly from the definition of rapidity in

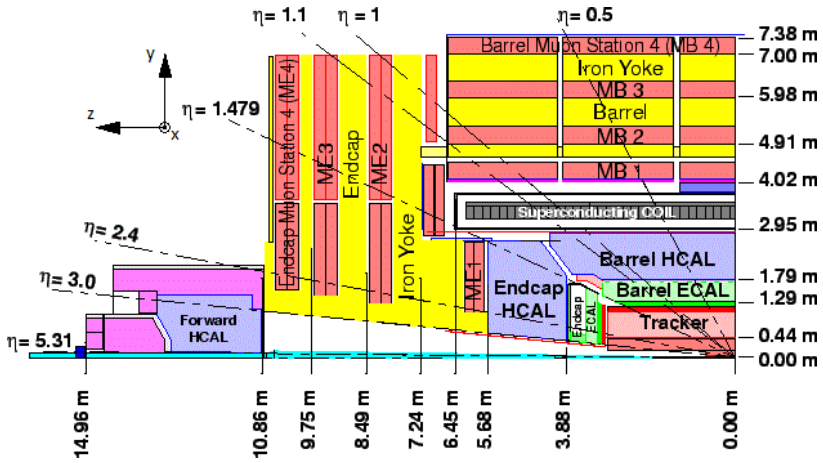

Figure 2.3 A slice of the CMS detector showing the coordinate system convention. The x, y, and z-axes are shown together with some typical η values.

special relativity as it quantifies only the boost in the z-axis. In the lab frame of particle colliders, the only boost comes from the momentum imbalance between colliding protons moving along the z-axis. As the rapidity difference Δy is

invariant under Lorentz transform, it is an important property for collider physics where measurements are done in the lab frame. Moreover in hadron colliders, the colliding constituent quarks and gluons carry different longitudinal momentum fractions x_L. This means that distributions measured as a function of Δy (and approximately for $\Delta\eta$) are not distorted by effects due to asymmetrical collisions between the particles and their constituents. Experimentally, η remains a more preferred variable over y as it is easier and faster to calculate to a good precision.

2.3.2 The CMS Trigger System

During runs, the beam bunches would cross up to 40 million times per second, with about a billion pp interactions per second at peak performance. Although this is advantageous for producing and detecting rare events, this poses a huge technical difficulty for the detector and computing system. Ideally, the data from all the channels in each sub-detector for every single event should be recorded. However, no computing hardware is able to handle such a write speed (40 MHz). Even assuming such a write speed can be achieved, the amount of data would reach 40 TB/s. To store such a large amount of data for a year of physics analyses would be unfeasible. Therefore the trigger system is developed to make careful choice of events which are worth keeping, and discarding the rest. Since most of the processes that go on in each pp collision are either easily produced in abundance or are already well understood, while interesting processes are much rarer, a mixed strategy for event sampling is required to ensure a good coverage of event samples for various physics analyses.

The CMS trigger is divided into two levels: Level 1 (L1) trigger, and High Level Trigger (HLT). The L1 trigger comprises of fast, programmable electronics installed within the detector itself, capable of making a decision within 3.2 µs. This quick decision can be made by using only data from the calorimeters and muon detectors. The L1 trigger reduces the data rate from 40 MHz to 100 kHz and passes data to the HLT, which is a software trigger that uses more than 13000 CPUs and all the available information from the various sub-detectors. This step cuts the data rate down to 100 Hz, keeping only samples of events that are interesting. Despite this, CMS still records several petabytes of data.

As part of the trigger system, two of the CMS detectors form an important role in telling the system when to start triggering. The Beam Pick-up Timing for

eXperiments (BPTX) is comprised of devices located 175 m from the IP on each side of the CMS detector. It provides the time and bunch structure information of the incoming beams with a time resolution of 0.2 ns. Using this information, it is possible to differentiate between cases where only a single beam bunch has crossed the detector to those with potential collisions from two bunches crossing on each side. The Beam Scintillator Counter (BSC) is another trigger system detector placed 10.86 m from the IP on both sides of the CMS detector. They detect particles emerging from the IP in the very forward (nearer the z-axis) direction within $3.23 < |\eta| < 4.65$. Being a forward detector, it picks up very soft particles that have small transverse momenta p_T. With a time resolution of 3 ns, it is able to reject events in which particles are detected by CMS but are produced by rare beam-gas collisions that happen in the LHC beampipes before the beam bunch reaches the IP. Both the BPTX and BSC are important detectors used to trigger zerobias and minimum bias data samples that will be used in the following analyses.

Chapter 3

Selected Topics in Soft QCD

3.1 Charged-particle Multiplicity Distribution

Charged-particle Multiplicity Distributions (MD) are one of the most basic yet powerful distributions that can be measured at particle colliders. The measurement of the charged-particle MD is made by simply counting the total number of charged-particles produced by a single pp collision. It aims to quantify the global characteristics of a pp collision by making inclusive measurements. Theoretically the MD should be measured in full phase space (in η and ϕ), i.e. any product of the process $pp \to X$, where X refers to any number of final-state particles, including neutrals. By its inclusiveness, the MD serves as a powerful measurement which differentiates between various model predictions.

Since most particle productions are dominated by soft processes, the particles captured in measurements of the MD have a large contribution from soft QCD. Charged-particle MD thus serve as a good tool to differentiate among various phenomenological models of soft QCD particle production.

Due to detector capabilities and constrains, generally only charged-particles are counted as the trackers (Sec. 2.3.1) which provide the spatial resolution required for high efficiency individual particle detection are only able to detect charged-particles. As most final-state particles are charged, this simplification does not change much of the physics. Thus henceforth, the MD will refer to charged-particle multiplicity distributions for simplicity.

Measurements of the MD are separated in windows (intervals) of pseudorapidity $|\eta|$ in order to distinguish between production processes that are kinematical

different. Results from the CMS Collaboration [9] are generally presented for $|\eta| = 0.5$, 1.0, 1.5, 2.0 and 2.4, as shown in figure 3.1 (left). The most inclusive measurement corresponds to counting only charged-particles within $|\eta| < 2.4$. This maximum is due to geometrical constrains of the detector and the fact that proton beams are injected in the z-axis.

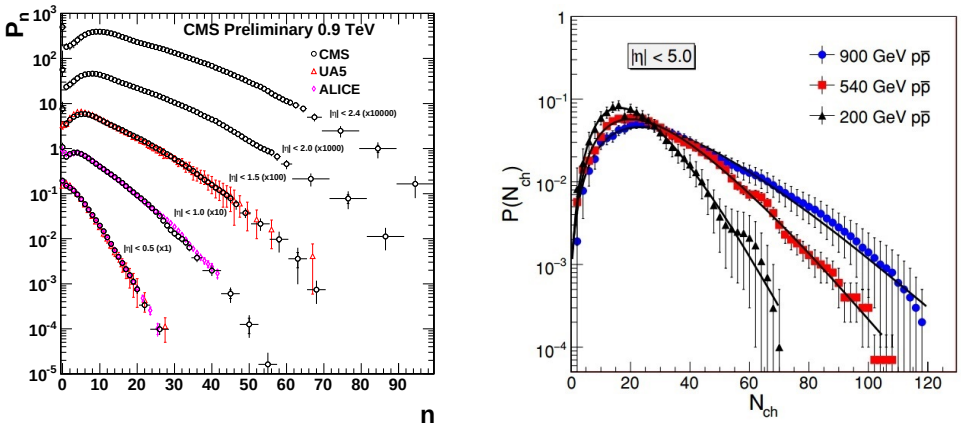

Figure 3.1 (Left) Charged-particle multiplicity distribution for pp collisions at $\sqrt{s} = 0.9$ TeV across all rapidity windows as measured by the CMS Collaboration [9]. The results are compared to the same measurements done by the UA5 Collaboration [10] for $p\overline{p}$ collisions and by the ALICE Collaboration [11] for pp collisions. (Right) Comparison of the charged-particle multiplicity distributions for $p\overline{p}$ collisions at $\sqrt{s} = 200$, 540, 900 GeV, and a phenomenological description by the NBD [12]. Data points are measured by the UA5 Collaboration [13, 10].

An important feature of the MD is that the shape of the distribution reveals information on the correlation and dynamics of particle production. If each detected particle from a single pp collision is produced independently, the MD would follow a Poisson distribution (PD). In the case where the particles decay and give rise to a cascade or shower, the resulting distribution would broaden to that of a Negative Binomial Distribution (NBD) [26–28].

Measurements of the MD have shown that the shape of the NBD provides a good description of experimental data [11, 12, 26–30]. However, this is broken by the appearance of a shoulder-like structure which becomes more prominent at higher centre-of-mass energies $\sqrt{s}$ [9, 11, 12, 29, 30]. Figure 3.1 (right) shows the comparison between the NBD [12] and the MD data for $p\overline{p}$ collisions at $\sqrt{s} = 200$, 540, and 900 GeV measured by the UA5 Collaboration [10, 13]. The NBD tends to describe poorly at larger $\sqrt{s}$ as the shoulder structure becomes more obvious at a multiplicity of around 60. The fit underestimates the data

points around the shoulder region and overestimates at larger multiplicities. This trend is corroborated by the ill fitting of the NBD at LHC energies [11, 29].

The failure of the NBD to describe MD data hints at the emergence of new particle production processes at higher $\sqrt{s}$. In order to understand these processes, various other distributions have been proposed. These include mixed models like two or three component [31, 32] models derived by means of weighted superposition of the NBD with other distributions, quantum optical inspired distributions [33, 34], parton branching models obtained from QCD and related models [35–37].

Distributions obtained from parton branching models include the NBD in the simplest case, and the more general Generalised Multiplicities Distribution (GMD) [35, 36] which describes a cascade of quarks and gluons which eventually hadronises into the observed final-state hadrons. Given m quarks and k' gluons produced from a collision, the GMD is given by

$$P_{GMD}(n;p,k,k') = \frac{\Gamma(n+k)}{\Gamma(n-k'+1)\Gamma(k'+k)}(1-p)^{n-k'}(p)^{k'+k}, \tag{3.1}$$

where Γ is the gamma function, n is the number of final-state hadrons, $k = m\tilde{A}/A$, and $p = \exp(-At) = (k'+k)/(\overline{n}+k)$. $A\Delta t$ and $\tilde{A}\Delta t$ are the probabilities of quark bremsstrahlung and gluon fission within the infinitesimal interval of the QCD evolution parameter $t = \frac{6}{11N_c-2N_f}\ln\left[\frac{\ln(Q^2/\mu^2)}{\ln(Q_0^2/\mu^2)}\right]$. The average multiplicity of final state hadrons $\overline{n}$ is controlled by the initial quark and gluon numbers as well as the branching probability through $\overline{n} = \frac{k'+k}{p} - k$.

If there are initial state fluctuations in gluon production, the GMD will be modified. This is described by a Weighted GMD (WGMD) [37]

$$P_{WGMD}(n;x_1,...,x_r,p,k) = \sum_{k'=0}^{n} P(k';x_1,...,x_r) \times P_{GMD}(n;p,k,k'), \tag{3.2}$$

where each independent collision event emits k' number of gluons with probability $P(k';x_1,...,x_r)$. In the particular case of Poissonian initial state fluctuations, the distribution becomes

$$P_{PGMD}(n;\overline{k'},p,k) = \sum_{k'=0}^{n} \frac{\overline{k'}^{k'}\exp\left(-\overline{k'}\right)}{k'!} \times P_{GMD}(n;p,k,k'), \tag{3.3}$$

where $\overline{k'}$ is the average initial number of gluons.

The Poisson weighted GMD (PGMD) provides reasonable description of CMS MD data without adding any additional parameters to the GMD. Figure 3.2 (left) shows an improved description of CMS MD data by the PGMD when a comparison is made with the GMD. Further details can be found in [37].

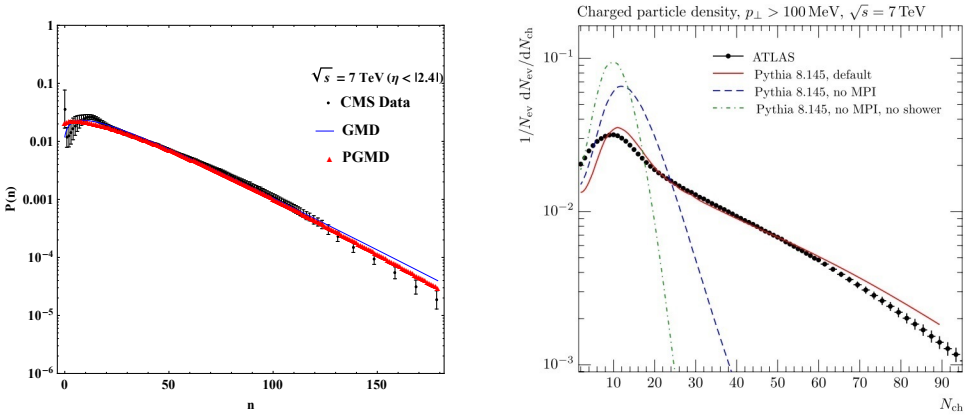

Figure 3.2 Comparison of various phenomenological models (left) GMD and WGMD, and (right) PYTHIA8 MC with and without MPI, against MD data at $\sqrt{s} = 7$ TeV.

Another physical mechanism that could explain the shoulder in MD is Multi-Parton Interaction (MPI) [38]. Since the proton is a composite object consisting of a partonic cloud, it is reasonable to expect multiple pairs of partons interacting within each pp scattering, as described in the MPI picture. As the number of partons increases with $\sqrt{s}$, MPIs are expected to play an increasingly important role. As seen in figure 1.10, the increase is almost completely in the low x_L regime dominated by the gluons. Thus most of the MPIs are expected to be soft.

The MPI mechanism is widely implemented in various modern MC generators to describe experimental data. The validity of MPI in describing MD data can be seen in figure 3.2 (right), where the MC generator PYTHIA8 without MPI fails to describe the data. With MPI implemented, the shoulder emerges and describes the data nicely.

An advantage of the MPI and its implementation in MC generators is that it should exhibit universality in describing collider data. The same collisions that result in the MD should also give rise to effects in other observables. Some of these include the underlying event and possibly long-range near-side angular correlations in pp collisions which will be discussed in the following sections.

3.2 Angular Correlations

In heavy ion collisions, two-particle angular correlations between final-state charged-particles provide evidence [39] for the presence of quark gluon plasma (QGP), a hot and dense fluid formed by the asymptotically free quarks and gluons of an energetic collision. Due to fluctuations in the initial collision geometry of the colliding nucleons, different pressure gradients develop which result in the hydrodynamical flow of the QGP in preferred directions. Of the various types of correlations that can be measured between two particles, the 2-dimensional (2D) $\Delta\eta$-$\Delta\phi$ correlation function is particularly interesting as it reveals much information on the physics of particle production including some novel effects.

The 2D correlation function, or per-trigger-particle associated yield, is quantified experimentally by matching each particle ("trigger") with every other remaining particle ("associated") in a given p_T range, obtaining the distribution

$$\frac{1}{N_{trig}}\frac{d^2N^{pair}}{d\Delta\eta d\Delta\phi} = B(0,0)\times\frac{S(\Delta\eta,\Delta\phi)}{B(\Delta\eta,\Delta\phi)} \tag{3.4}$$

of the number of matched pairs N^{pair} with relative angular differences in $\Delta\eta$ and $\Delta\phi$. The distribution is normalised by N_{trig}, the number of trigger particles per event, to get the correlated yield per trigger particle. The signal and background distributions $S(\Delta\eta,\Delta\phi)$ and $B(\Delta\eta,\Delta\phi)$ are defined as

$$S(\Delta\eta,\Delta\phi) = \frac{1}{N_{trig}}\frac{d^2N^{same}}{d\Delta\eta d\Delta\phi} \tag{3.5}$$

$$B(\Delta\eta,\Delta\phi) = \frac{1}{N_{trig}}\frac{d^2N^{mix}}{d\Delta\eta d\Delta\phi}, \tag{3.6}$$

where N^{same} and N^{mix} refer to the number of pairs in the same event and different events respectively. Random combinatorial background and pair acceptance effects are removed by dividing the signal by the background distribution.

Particle correlations due to various effects can be seen in the structure of the 2D correlation function, as shown in figure 3.3 (top) [40]. The correlation functions between *pPb* and *PbPb* collisions at $\sqrt{s} = 2.76$ and 5.02 TeV respectively are shown. Peak structures due to the production of jets are strongly reflected at $(\Delta\eta,\Delta\phi) = (0,0)$. Due to momentum conservation in di-jet events, a corresponding away side peak centred around $(0,\pi)$ can be seen, albeit with a

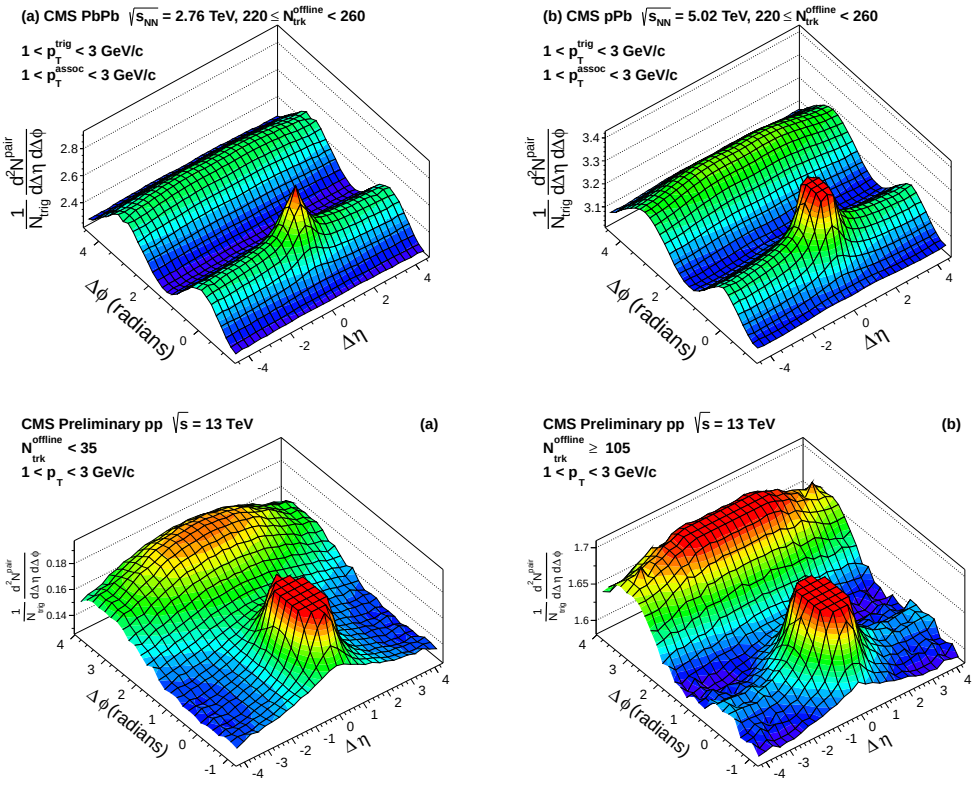

Figure 3.3 The 2D two-particle angular correlation functions for (top left) 2.76 TeV *PbPb* and (top right) 5.02 TeV *pPb* collisions using charged-particle pairs with $1 < p_T < 3$ GeV in events where the particle multiplicity is within $220 \leq N_{trk}^{offline} < 260$. The same distribution is shown for *pp* collisions at 13 TeV with (bottom left) $N_{trk}^{offline} < 35$ and (bottom right) $N_{trk}^{offline} \geq 105$. The right and bottom figures have sharp peaks that are truncated to emphasise the other correlation structures.

larger $\Delta\eta$ spread due to initial state fluctuations of partonic x_L in the colliding nucleons and hydrodynamical effects. What is particularly interesting is the long-range near-side structure, a ridge in the $\Delta\phi \approx 0$ side of extended reach in $\Delta\eta$. This ridge can be described by hydrodynamical models of the collective flow of strongly interacting QGP medium with fluctuations in the particle production region's energy density [41–46].

Similar long-range near-side ridge structures were first observed in high multiplicity events for pp collisions at $\sqrt{s} = 7$ TeV [47] and pPb collisions at $\sqrt{s} = 5.02$ TeV [48, 40], and confirmed again for pp collisions at $\sqrt{s} = 13$ TeV [49]. Figure 3.3 shows the results at 13 TeV in low multiplicity ($N_{trk}^{offline} < 35$) events (bottom left) as opposed to high multiplicity ($N_{trk}^{offline} \geq 105$) events (bottom right).

This observation is most surprising for pp collisions since the partonic density is not expected to be high enough to form QGP. However this novel phenomenon indicates the emergence of a different means of particle production when the multiplicity of the event increases. Further insight can be gleaned by comparing the integrated associated yield for different collisions systems as a function of $N_{trk}^{offline}$ as shown in figure 3.4. The yield is compatible with zero up to $N_{trk}^{offline} \approx 40$, with an approximate linear rise after. Interestingly, the yield depends only on the collision system and not on $\sqrt{s}$.

Naturally, hydrodynamical models which were applied to heavy ion collisions [50–54] are now extended to describe the same phenomena for pp collisions. Various other production mechanisms have also been proposed to explain this phenomenon. These include colour connections in the longitudinal direction [55–58], interaction of the medium with jets [59], colour glass condensates [60–62], and even MPI [63, 64].

Although it remains to be seen whether MPI or other particle production mechanisms are at play in the description of the long-range near-side ridge in pp collisions, MPI have already been widely implemented in various Monte Carlo (MC) generators to explain the UE activity, which will be the subject of focus in the next chapter.

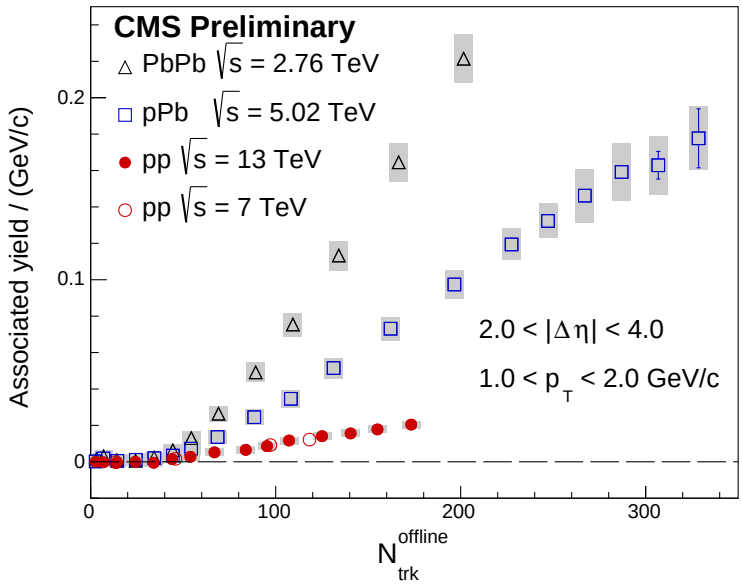

Figure 3.4 The integrated associated yield as a function of the multiplicity $N_{trk}^{offline}$ for the near-side of the correlation function for pp collisions at $\sqrt{s} = 7$ and 13 TeV, pPb collisions at $\sqrt{s_{NN}} = 5.02$ TeV, and $PbPb$ collisions at $\sqrt{s_{NN}} = 2.76$ TeV. Error bars denote statistical uncertainties while the shaded areas and boxes represent the systematic uncertainties.

Chapter 4

The Underlying Event

The measurement of the underlying event activity in pp collisions is performed with respect to the highest p_T^{jet} charged-particle jets at $\sqrt{s} = 2.76$ TeV [65–67] and highest p_T/p_T^{jet} charged-particles/charged-particle jets at $\sqrt{s} = 13$ TeV [68, 69]. The analyses uses data collected from the CMS experiment at the LHC as well as various Monte Carlo generated samples for comparison. The results and analyses in this chapter constitute work done at the LHC, CERN.

4.1 Introduction

The underlying event (UE) activity in a high-energy hadron-hadron collision is defined as the radiation that is not part of the final-state activity originating from the most energetic parton scattering of the collision. Unlike e^+e^- collisions where the colliding particles have no substructure, hadron production in a pp collision is understood to originate from multiple scatterings of the partonic constituents in a collision region of extended spatial dimensions. Theoretically, a large portion of the partons will not participate in the scattering and will be emitted in the very forward direction and should also contribute to the final-state radiation. However most will not be observed due to physical constrains in the detector design. Partons produced from the collision reduce their virtuality through gluon radiation and quark-antiquark splittings, and finally fragment into hadrons at scales approaching 0.2 GeV (Λ_{QCD}). Ideally, the produced hadrons in an event are identified as those coming directly from the fragmentation of partons resulting from the scattering with the largest momentum transfer (hard scattering),

and the rest which forms the UE. The UE activity is then comprised of hadrons coming from (A) initial- and final-state radiation (ISR, FSR) from the hard scattering (B) softer partonic scatters in the same pp collision (multiple parton interactions, or MPI) possibly with their own initial- and final-state radiation, and (C) proton remnants concentrated along the beam direction. In reality, the final-state hadrons cannot be uniquely identified as coming from any one of these classifications. However, experimentally "clean" observables that are relatively uncontaminated by the radiation from the hardest partonic scattering can be defined for the UE activity.

An accurate understanding of the UE is required for precise measurements of standard model processes at high energies and searches for new physics. The UE affects measurements of isolated high transverse momentum p_T leptons or photons; and because it dominates most of the hadronic activity from the additional pp collisions taking place in a given bunch crossing (pileup) at the high luminosities achieved by the CERN LHC, it is important to understand the impact of its activity. Since the semi-hard and low-momentum partonic processes, which dominate the UE cannot be adequately calculated with pQCD methods alone, the UE activity requires a phenomenological description containing parameters that must be tuned to data.

The experimental observables sensitive to the UE activity can be defined by exploiting the topological structure of pp collisions with a hard scattering. One example is the study of particle properties in regions away from the direction of the products of the hard scattering. At the Tevatron, the CDF experiment measured UE observables using inclusive jet and Drell–Yan (DY) events in $p\bar{p}$ collisions at the centre-of-mass energies of $\sqrt{s} = 0.63$, 1.8, and 1.96 TeV [70–72]. In pp collisions at the LHC, the ALICE, ATLAS, and CMS experiments have carried out UE measurements at $\sqrt{s} = 0.9$ and 7 TeV using events containing a leading (highest p_T) charged-particle jet [14, 15, 73], a leading charged particle [74, 75], a DY lepton pair [76], or a $t\bar{t}$ di-jet pair [77]. In this chapter, we study the UE activity in pp collisions at $\sqrt{s} = 13$ TeV by measuring the average multiplicity and scalar transverse momentum sum (Σp_T) densities of charged particles in the azimuthal region orthogonal to the direction of the leading charged-particle or jet, referred to as the transverse region.

At a given centre-of-mass energy, the UE activity is expected to increase with the momentum transfer between the interacting partons (hard scale), probed by

the p_T of the leading charged-particle/jet of the event. On average, increasingly hard partonic interactions result from pp collisions with decreasing impact parameters between the two protons, which in turn enhance the overall hadronic activity originating from MPI until a saturation is reached for central collisions with maximum overlap [78, 79]. At the same time, the activity related to the ISR and FSR components also increases with the hard scale. For events with the same hard scale, probed by the p_T of jets or DY pairs, the MPI activity rises with $\sqrt{s}$ as more partons are expected in the protons at increasingly smaller parton fractional momenta $x_L \sim 2\,p_T/\sqrt{s}$ [78, 79]. Hence, studying the UE as a function of the hard scale (p_T/p_T^{jet}) at several centre-of-mass energies provides an insight into the UE dynamics and its evolution with the collision energy, further constraining the MPI model parameters.

4.2 Methodology and Observables

The physical variables defined to describe the intensity of the UE activity are the average multiplicity density $\langle N_{ch}\rangle / [\Delta\eta\Delta(\Delta\phi)]$ and scalar p_T sum density $\langle \sum p_T\rangle / [\Delta\eta\Delta(\Delta\phi)]$ of the charged-particles, which are studied as a function of the leading p_T/p_T^{jet}. The densities are obtained by dividing the event averaged charged-particle multiplicity and Σp_T by the solid angle in $\eta \times \phi$ space where the particles are detected. This is given by $60° < |\Delta\phi| < 120°$ (i.e. $2\pi/3$) and $|\eta| < 2$ (i.e. 4), giving a solid angle of $8\pi/3$.

In order to separate observables sensitive to the UE activity from the radiation of the hardest partonic scattering of the event, the topological structure of an event can be used. By convention, the direction of the leading charged-particle/jet is used as a reference to define three angular regions in ϕ as follows:

- "Towards" region defined by $|\Delta\phi| < 60°$,
- "Transverse" region defined by $60° < |\Delta\phi| < 120°$,
- "Away" region defined by $|\Delta\phi| > 120°$,

where $\Delta\phi$ is the relative azimuthal angle between a selected charged-particle and the direction of the leading charged-particle/jet. This is illustrated in figure 4.1.

As shown in previous results [70–72], the activities in the towards and away regions are dominated by initial- and final-state radiations and mostly sensitive

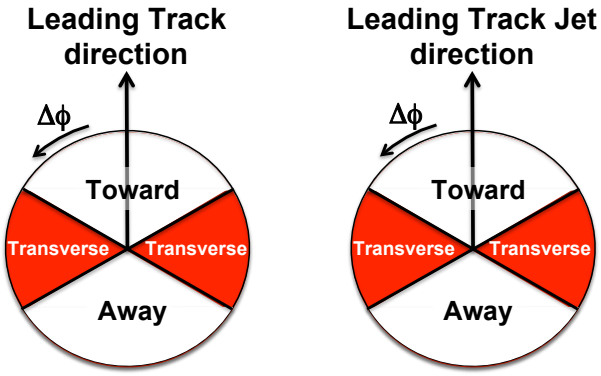

Figure 4.1 Illustration of correlations in azimuthal angle $\delta\phi$ relative to (*l*eft) the direction of the leading track, or to (right) the leading track-jet.

to the $2 \rightarrow 2$ hard scattering which results in jets and high p_T particles. For jet events (e.g. Figs. 1.4 and 1.5: di-jet or tri-jet topology), the interaction scale of the collision is probed by the highest p_T (leading) particle or jet. The flow of particles and energy in the direction of the jet reflect dynamics of QCD physics, calculable by pQCD. By looking away from regions dominated by jet radiation, the $\langle N_{ch} \rangle / [\Delta\eta\Delta(\Delta\phi)]$ and $\langle \sum p_T \rangle / [\Delta\eta\Delta(\Delta\phi)]$ densities can be measured for the soft radiative component. This can be illustrated by looking at the densities as a function of the relative azimuthal angle $\Delta\phi$ between the each particle and the leading particle (Fig. 4.2). The $\langle \sum p_T \rangle / [\Delta\eta\Delta(\Delta\phi)]$ density reveals significant contribution from the di-jet in the towards and away regions while having a small but non-zero contribution from softer radiation density in the transverse region. Therefore in the following analyses, we will only focus on the transverse region which is most sensitive to the UE activity.

The transverse region can be further separated into two sides, which together with their combinations, are able to yield 4 types of densities (for both $\langle N_{ch} \rangle / [\Delta\eta\Delta(\Delta\phi)]$ and $\langle \sum p_T \rangle / [\Delta\eta\Delta(\Delta\phi)]$):

- Transverse/transAVE densities: The event averaged densities in the transverse region.
- TransMAX densities: The event averaged densities on the half of the transverse region with higher multiplicity or Σp_T.
- TransMIN densities: The event averaged densities on the half of the transverse region with lower multiplicity or Σp_T.

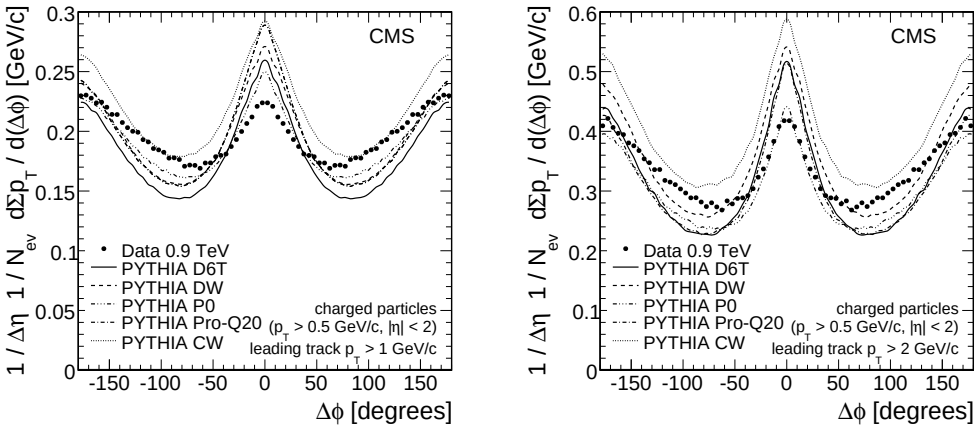

Figure 4.2 Average scalar sum p_T density of charged particles as a function the azimuthal angle difference $\Delta\phi$ between each particle relative to the leading particle with (left) $p_T > 1$ GeV and (right) $p_T > 2$ GeV. The data points (filled circles) are compared to several PYTHIA MC tunes (lines) which went through full detector simulation.

- TransDIF densities: The difference between TransMAX and TransMIN densities.

This is illustrated in figure 4.3. The method was first suggested by Bryan Webber [80] and implemented by Jon Pumplin [81] but never done at LHC energies.

For events with large initial- and final-state radiation, the transMAX densities include contribution from the third jet while both transMAX and transMIN densities receive contributions from MPI and beam-beam remnants. The transMIN activity is therefore sensitive to MPI and beam-beam remnants while the transDIF activity (transMAX minus transMIN) is sensitive to initial- and final-state radiation. The measurement of the transverse UE densities are thus able to isolate the leading contribution of $2 \to 2$ scattering from $\langle N_{ch} \rangle / [\Delta\eta\Delta(\Delta\phi)]$ and $\langle \sum p_T \rangle / [\Delta\eta\Delta(\Delta\phi)]$, and with further separation into transMAX, transMIN, and transDIF densities, the sensitivity of the observables to soft and hard QCD effects should become more apparent in the results.

As the detector makes measurements of the particles that pass through its various subsystems, it also affects them slightly due to interactions with the detector system and its inefficiencies. While the main objects of interest that can be studied in this analysis are the tracks left by the particles as they move through the silicon tracker system (subsection 2.3.1), a useful measurement of the UE activity requires an extraction of the "true" level of activity at the particle level

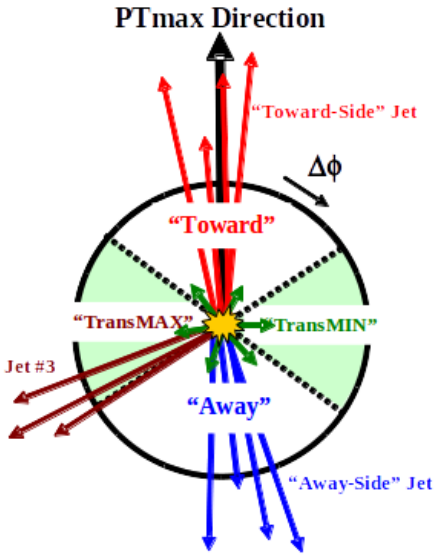

Figure 4.3 Illustration of the separation of the overall transverse region into the transMAX and transMIN regions based on the activity in the region.

(i.e. measurement by a perfect detector). Thus the track level UE observables have to be corrected for detector effects.

As such, the subsequent sections will be focused on the exposition of two separate analyses (UE at 2.76 and 13 TeV) that extract particle level results with the following methodology:

1. **Data and Monte Carlo event samples**: The CMS data samples are recorded only if it is flagged by the trigger system discussed in subsection 2.3.2. Based on information from the BPTX and BSC detectors, two different trigger streams can be obtained: zerobias and minimum bias.

2. **Event and track selection**: Once the triggered/flagged events have been selected, additional event selection criteria are made to obtain a subset of good quality events. Cuts on the track collection also have to be made to ensure selection of high quality tracks.

3. **Data correction**: Track observables are then corrected to particle level observables which can then be compared to various MC generator predictions.

4. **Evaluation of systematic uncertainties**: Systematic uncertainties coming from various sources are evaluated and propagated through the correction algorithm in order to obtain uncertainties at the particle level.

4.3 Measurement of the UE at 2.76 TeV

A measurement of the UE activity in pp collisions is performed using events with charged-particle jets produced in the central pseudorapidity region ($|\eta^{\text{jet}}| < 2$) and with transverse momentum $1 \leq p_T^{jet} < 100$ GeV. The analysis uses a data sample collected $\sqrt{s} = 2.76$ TeV [65–67] with the CMS experiment at the LHC. The UE densities are measured as a function of p_T^{jet} using charged particles with $|\eta| < 2$ and $p_{\text{T}} > 0.5$ GeV, in the transverse region with respect to the leading jet. The transMAX, transMIN, and transDIF densities are measured for the first time at the LHC, separating the various components of the UE activity. The measurements are compared to previous results at 0.9 and 7 TeV, and to predictions of several Monte Carlo event generators, providing constraints on the modelling of the UE dynamics. This allows the CMS Collaboration to understand soft QCD radiation better and in turn make more precise measurements of in searches of standard model and new physics processes.

4.3.1 Data and Monte Carlo samples

Data Samples

The analysis data set forms a subset of all pp collision data collected by the CMS Collaboration at $\sqrt{s} = 2.76$ TeV, where the low PU conditions are suitable for the UE analysis. Pileup occurs in 6.2% of the events, giving an average of 0.12 pp collisions per event. The data set corresponds to an integrated luminosity of 0.30 nb^{-1}.

The minimum bias trigger used provides minimal amount of bias to the triggered collision data while improving the effectiveness of collecting useful data. It requires activity in both the BSC counters in coincidence with signals from both BPTX devices. This means that at least a single track originating from the IP is detected in the forward (large $|z|$) region. While this ensures that the collected data contains pp collisions(s) (as opposed to detection of cosmic

rays or other background sources), it inevitably introduces a small bias on the event selection. However, the selection bias is known to have negligible effect on the UE measurements [82]. On the other hand, if every event were to be recorded, the relatively low instantaneous luminosities of the beams translate to having no pp interaction in many of the bunch crossings, wasting memory space and processing time. Thus the minimum bias trigger serves to improve the effectiveness of data collection by flagging only events with a collision(s) for recording.

In order to enhance event samples at high jet p_T, additional single jet triggers are also used. As the inclusive jet cross-section dies out at large p_T, the event statistics within the minimum biased sample drops correspondingly. The jet triggers are designed to enhance the collection of events with high p_T jets by flagging an event for recording if the transverse energy E_T recorded by the ECAL or HCAL exceeds a selected set of thresholds.

Single jet triggers with thresholds of 20 GeV (Jet20) and 40 GeV (Jet40) are used in combination with the minimum bias trigger. The triggered samples are combined by finding an offline p_T^{jet} threshold at which a transition can be made to the single jet triggered sample without bias. This is illustrated in figure 4.4, where events flagged by the minimum bias and Jet20 trigger are ordered in p_T^{jet}. As the minimum bias sample size dies out at higher p_T^{jet}, the Jet20 trigger starts to flag more events which are also triggered by the minimum bias trigger. At even larger p_T^{jet}, the Jet20 sample completely dominates while the minimum bias sample dies out. The transition point in p_T^{jet} where the co-triggered event is 100% of minimum bias sample is the threshold at which Jet20 triggered events are no longer biased.

The threshold is decided by evaluating the trigger efficiency curve as a function of jet p_T. The p_T^{jet} value, where efficiency reaches to $\sim$100%, is considered for offline selection i.e. for Jet20 and Jet40 triggers, an offline cut of 25 GeV and 50 GeV, respectively, is applied. This selection can appreciated by looking at figure 4.5 which shows the (top left) leading jet p_T distribution, (top right) efficiency of Jet20 trigger w.r.t. minimum bias trigger (fitted with a constant from 25 GeV to 50 GeV), (bottom left) ratio of all Jet20 triggered events w.r.t minimum bias triggered events (fitted with a turn on curve from 15 GeV to 50 GeV) and (bottom right) efficiency of Jet40 trigger w.r.t. Jet20 trigger (fitted with a turn on curve from 25 GeV to 100 GeV). In figure 4.5 (top left), events

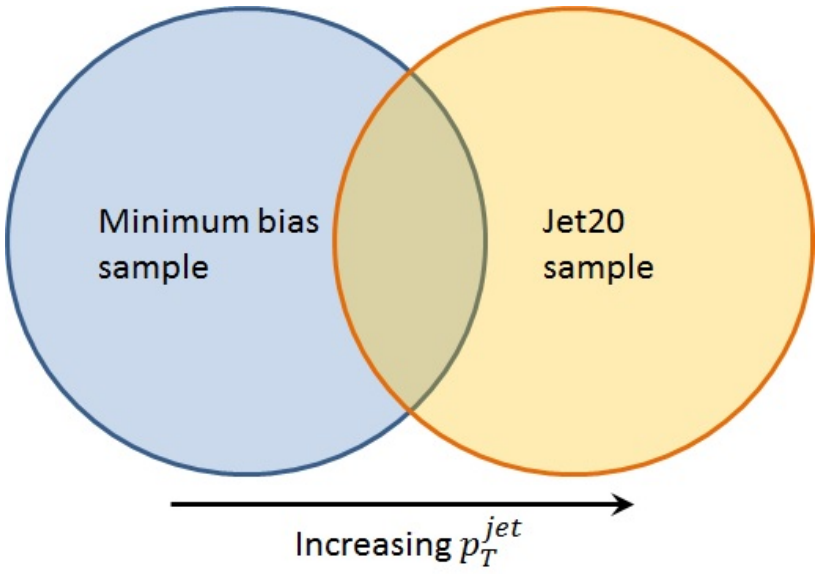

Figure 4.4 Venn diagram showing the set of minimum bias and Jet20 triggered events ordered in p_T^{jet}. The overlapping region is the sample of events where both triggers fired. The p_T^{jet} for which the co-triggered event is 100% that of the minimum bias event sample within the p_T^{jet} slice is set as the threshold.

are weighted by the trigger prescales (PRSC), a pre-selected scale that records only 1/PRSC of the triggered events if the post-trigger rate is still too high for the computer system to record everything.

The efficiency of the Jet20 trigger w.r.t. the minimum bias trigger can be appreciated by looking at figure 4.5 (top right) where the efficiency reaches a maximum of around 100% (from the constant fit) at 25 GeV. Due to the lack of statistics causing huge error bars in the efficiency above 20 GeV, it is not possible to get a reasonable fit of a turn on curve on figure 4.5(top right) and only a constant fit is performed. Instead, a comparison to the efficiency can be made with the ratio between all Jet20 events against minimum bias events (Fig. 4.5(bottom left)) where a turn on curve is fitted. The plateau of the turn on curve shows compatibility with the constant fit in the efficiency plot within error bars. Similarly from figure 4.5 (bottom right), we can see that the efficiency of Jet40 triggered events w.r.t. Jet20 trigger reaches a maximum of almost 100% at 50 GeV.

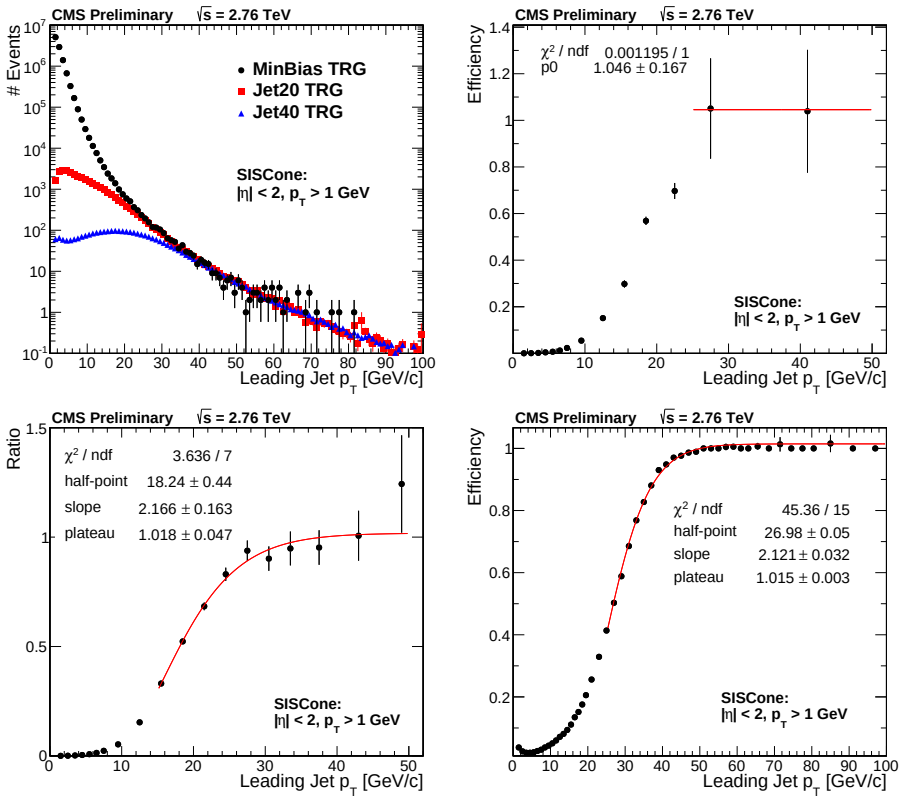

Figure 4.5 (Top left) Leading jet p_T, (top right) efficiency $\left(PRSC\_Jet20 \times \frac{N_{MinBias+Jet20}}{N_{MinBias}}\right)$ of Jet20 trigger w.r.t. minimum bias trigger, (bottom left) ratio $\left(\frac{PRSC\_Jet20 \times N_{Jet20}}{PRSC\_MB \times N_{MinBias}}\right)$ of all Jet20 triggered events w.r.t minimum bias triggered events and (bottom right) efficiency $\left(PRSC\_Jet40 \times \frac{N_{Jet20+Jet40}}{N_{Jet20}}\right)$ of Jet40 trigger w.r.t. Jet20 trigger. In the top left plot, events are weighted according to trigger prescales.

Monte Carlo Samples

Monte Carlo generated events are used to correct the data for detector effects and for comparisons with the final unfolded data. Various MC generators and tunes of each generator are available, differing in their physics modelling as well as their parameter values. In this analysis, the PYTHIA [83, 84] and HERWIG++ [85] generators are used, with various tunes that are described below.

PYTHIA6 and PYTHIA8 are general purpose generators that use the Lund string hadronisation model [86]. In PYTHIA6 there is an option of virtuality-ordered or p_T-ordered parton shower, with the latter used in most recent tunes. In PYTHIA8, only the p_T-ordered parton shower is used. Hadron-hadron interactions are described by splitting the total inelastic cross-section into non-diffractive processes, dominated by t-channel (Mandelstam variable) gluon exchange, and diffractive processes involving a colour-singlet exchange. The diffractive processes are further divided into single-diffractive dissociation, where one of the initial hadrons remains intact while the other dissociates, and double-diffractive dissociation where both hadrons dissociate. Such events tend to have large gaps in particle production at central rapidity. Central diffraction, in which both hadrons remain intact and particles are produced in the central region, is neglected as it is rare.

The $2 \rightarrow 2$ non-diffractive processes, including MPI, are described by lowest-order pQCD calculations with the divergence of the cross-section as $p_T \rightarrow 0$ regulated with a phenomenological model. Various tunable parameters control the behaviour of this regularisation as well as the matter distribution of partons within the hadrons and colour reconnection. When p_T-ordered parton showers are used, the MPI and parton shower are interleaved in one common sequence of decreasing p_T values. For PYTHIA6 the interleaving is between the initial-state shower and MPI only, while for PYTHIA8 it also includes final-state showers. Since the p_T-ordered showers and interleaving with MPI are considered to be a model improvement, the most recent PYTHIA6 tunes are made with this configuration. This is also the only configuration available in PYTHIA8.

A pomeron-based approach is used to describe diffractive events, using the Schuler and Sjöstrand [87] parametrisation of the pomeron flux. In PYTHIA6 the diffractive dissociations are treated using the Lund string model, producing final-state particles with limited p_T. In PYTHIA8 the dissociations are treated

like this only for events with a diffractive system with a very low mass; in higher mass systems, diffractive parton distributions from H1 [88] are used to include diffractive final-states, which are characteristic of hard partonic interactions. In this case, the full machinery of MPI and parton showers is used and yields a significantly harder p_T spectrum for final-state particles.

HERWIG++ uses an angular-ordered parton shower and the cluster hadronisation model [85]. It has an MPI model similar to the one used by the PYTHIA generators, with tunable parameters for regularising the behaviour at very low momentum transfer, but does not include the interleaving with the parton showers. Hadron-hadron collisions are simulated by applying the MPI model even in events with no hard scattering. It is therefore possible to generate an event with zero $2 \to 2$ partonic scatters, in which only beam remnants are produced, with nothing in between them. While HERWIG++ has no explicit model for diffractive processes, these zero-scatter events will look similar to double-diffractive dissociation.

In these MC generator models, MPIs play a central role in the description of UE activity, where the MPI activity is expected to increase with $\sqrt{s}$ [78, 79]. Data taken at the Tevatron [70–72] and, more recently, at the LHC for different centre-of-mass energies [14, 15, 74, 75] have provided a rich sample of experimental input which has focused efforts towards the development of how to model these interactions in MC generators.

PYTHIA6 and PYTHIA8 model the energy dependence of MPI as an exponential function of the centre-of-mass energy with tunable parameters[83, 84]. Explicitly, the model relies on two fundamental assumptions [89]:

- The ratio of the $2 \to 2$ partonic cross section, integrated above a p_T cutoff scale, to the total inelastic pp cross section is used as a measure of the amount of MPI. A factor with a free parameter, p_{T0}, is introduced to regularise an otherwise divergent differential partonic cross section,

$$\frac{d\sigma}{dp_T{}^2} \propto \frac{\alpha_s^2\left(p_T^2\right)}{p_T^4} \to \frac{\alpha_s^2\left(p_T{}_0{}^2 + p_T^2\right)}{\left(p_T{}_0{}^2 + p_T^2\right)^2}, \tag{4.1}$$

  with

$$p_{T0}(\sqrt{s}) = p_T{}_0^{REF}\left(\frac{\sqrt{s}}{\sqrt{s_0}}\right)^{\varepsilon}. \tag{4.2}$$

The parameters p_{T0}^{REF} and ε characterise the energy dependence of p_{T0}, with p_{T0}^{REF} being the value of p_{T0} at the reference centre-of-mass energy of $\sqrt{s_0} = 1.8$ TeV.

- A Poisson distribution for the number of MPI in an event, with a mean that depends on the overlap of the matter distribution of the hadrons in impact parameter space.

HERWIG++ also assumes that MPI varies with the colliding energy and has recently automated this process in more recent versions (newer than 2.6). The energy dependence of MPI in HERWIG++ is also an exponential function of the centre-of-mass energy with tunable parameters [85].

In this analysis, several event generator tunes are used. These are the PYTHIA6 [83] tune Z2/Z2* [90] and tune CUETP6S1 [91], PYTHIA8 [84] tune 4C [92] and CUETP8S1 [91], and HERWIG++ tune UE-EE-5C [85, 93]. These event generators and tunes differ in the treatment of initial- and final-state radiation, hadronisation, and in the choice of MPI parameters, colour reconnections, and cutoff values for the MPI mechanism. The parameter values were obtained by optimising the comparison between predictions and UE data, especially those measured with LHC pp collision. However, minimum bias observables and data from the Tevatron, collected at lower centre-of-mass energies, were also used to obtain some of these tunes.

The Z1 tune [94] of PYTHIA6 (predecessor of Z2 and Z2*) adopts the results of a global tuning performed by the ATLAS Collaboration [95] and uses the fragmentation and colour reconnection parameters of the ATLAS AMBT1 tune [96, 97]. The Z1 tune uses the CTEQ5L PDF and its parameters related to the MPI regularization cutoff and its energy dependence are adjusted to describe previous CMS measurements of the UE activity in hadronic events [15]. The Z2 tune of PYTHIA6 is then updated to use CTEQ6L1 [98] and with an optimal value of $p_{T0}^{REF} = 1.832$ GeV. The centre-of-mass energy dependence $\varepsilon = 0.275$, as with the Z1 tune. PYTHIA6 Z2* tune is derived from the Z2 tune by tuning the value of p_{T0}^{REF} and ε with the automated "PROFESSOR" tool [99], yielding $p_{T0}^{REF} = 1.921$ and $\varepsilon = 0.227$.

PYTHIA8 tune 4C [92] is tuned by making comparisons to minimum bias and UE data from the LHC following the procedure of the ATLAS AMBT1 tune [97], but includes ALICE minimum bias multiplicity data as well. The

Table 4.1 Summary of the parameters of the Monte Carlo generator tunes.

Tune	$p_{T_0}^{REF}$ (GeV)	$\sqrt{s_0}$ (GeV)	ε
Z2	1.832	1800	0.275
Z2*	1.921	1800	0.227
CUETP6S1	1.9096	1800	0.2479
4C	2.085	1800	0.19
CUETP8S1	2.1006	1800	0.2106
UE-EE-5C	3.91	7000	0.33

values of the $p_{T_0}^{REF}$ and ε parameters for the 4C tune are 2.085 GeV and 0.19, respectively. The effective value of p_{T0} is about 2.76 GeV at $\sqrt{s} = 7$ TeV for both Z2 and 4C tune.

PYTHIA6 CUETP6S1 and PYTHIA8 CUETP8S1 [91] are two UE tunes obtained by simultaneously fitting CDF data from $\bar{p}p$ collisions at 300 GeV, 900 GeV, and 1.96 TeV together with CMS data for pp collisions at 7 TeV. The UE model parameters were constrained for a wide range of centre-of-mass energies, which is expected to allow for more precise predictions at 13 TeV and 14 TeV.

The HERWIG++ generator tune UE-EE-5C [93] is tuned to LHC data and includes important final-state effects due to colour reconnections, and an exponential centre-of-mass energy dependence of MPI. The parameters used are $p_{T_0}^{REF} = 3.91$ and $\varepsilon = 0.33$, where $p_{T_0}^{REF}$ is defined at the centre-of-mass energy of 7 TeV.

A summary of the differences in the values of the parameters between the various MC generator tunes is given in table 4.1. All MC generator used in this analysis adopts the CTEQ6L1 [98] PDF.

Theoretically, MC generated events can always be generated such that there will be sufficient statistics at large p_T^{jet}. However to save processing time that would otherwise be used to wait for sufficient statistics to be populated at large p_T^{jet}, di-jet events are specifically generated with $\hat{p}_T$ larger than 15 and 50 GeV. Minimum bias and di-jet MC events are then merged according to their $\hat{p_T}$ values.

In order to compare the performance of the MC event generators with respect to data, detector response was simulated in detail using the GEANT4 package [100] and events were processed and reconstructed in the same manner as colli-

sion data. Pileup is not included in MC, as the level of pileups in collision data is small. They are quantified and incorporated as systematics in section 4.3.4.

4.3.2 Event and track selection

In this section we will discuss the event and track selection criteria following the same method as previous UE measurements [15, 101]. These selection criteria are briefly summarised below.

In order to remove beam halo, a beam scraping filter is applied to the triggered events. Event selection further requires exactly one primary vertex candidate. Additional quality criteria on the selected vertex candidate apply, following the prescriptions:

- the degree of freedom ($n\,dof$) has to be greater than 4, and it must be well centered with respect to the z-position of the beam-spot within ± 10 cm.

This event selection procedure contributes a bias of under 5% at the lowest leading track-jet p_T and drops to and average of 3% at higher p_T. Correction for the event selection is discussed in section 4.3.3 and the bias of the choice of degree of freedom on the correction is investigated in section 4.3.4.

For each selected event the reconstructed track collection needs to be cleaned up from undesired tracks, namely secondaries from the decay of long lifetime particles and background (e.g. combinatorial background and beam halo associated tracks). Fake tracks coming from mis-reconstruction are removed by requiring tracks to pass a selection criteria based on the minimum fraction of consecutive hits in the silicon tracker layers, depending on the pseudorapidity region. Secondaries are removed by requiring the impact parameter significance d_0/σ_{d_0} and significance of z separation between the track and primary vertex d_z/σ_{d_z} each to be less than 3. In order to remove tracks with poor momentum measurement, we require the relative uncertainty of the momentum measurement $\sigma(p_T)/p_T$ to be less than 5%. Selected tracks have $p_T > 0.5$ GeV and $|\eta| < 2.0$ to ensure good reconstruction efficiency. Figure 4.6 shows the comparison of different track properties (top left) p_T, (top right) η, (centre left) ϕ, (centre right) d_0/σ_{d_0}, (bottom left) d_z/σ_{d_z}, and (bottom right) σ_{p_T}/p_T for data and MC predictions. There is reasonable agreement between data and MC predictions and level of agreement is same as that observed in 7 TeV measurement. There is somewhat

events	exactly one vertex	NDOF>4, $\lvert z^{vtx} - z^{beamspot}\rvert < 10$ cm
	leading track-jet	$p_T > 1$ GeV, $\lvert\eta\rvert < 2.0$
tracks	$p_T > 0.5$ GeV, $\lvert\eta\rvert < 2.0$ $\lvert d_0/\sigma_{d_0}\rvert < 3.0$, $\lvert d_z/\sigma_{d_z}\rvert < 3.0$, $\sigma_{p_T}/p_T < 0.05$	

Table 4.2 Summary of event and track selection at detector level.

larger deviation for d_z/σ_{d_z} which is due to higher pileup contribution that gives rise to a greater contribution of tracks coming from other vertices. Table 4.2 summarises the event and track selection criteria applied at detector level.

Track reconstruction efficiency and fake rate are calculated by matching reconstructed charged particles with generated charged particles within a cone of radius parameter of 0.05. Average reconstruction efficiency for the selected tracks is about 85% which drops to 75% for low p_T tracks and large η as can be concluded from figure 4.7 (left column). Figure 4.7 (right column) shows the fake rate of selected tracks as a function of track p_T, η, and ϕ. It is clear that the average fake rate is about 2% but increases if tracks have small p_T or lie at large η.

The event scale and reference direction for identification of the UE sensitive region are defined using leading track-jets [102] (jets reconstructed from the clustering of tracks). The track-jet collection used in this study consist of jets that are reconstructed from charged tracks instead of calorimeter towers. Track-jets are reconstructed using the SisCone [103] algorithm with a distance parameter of 0.5. In order to have a valid comparison with previous UE measurements, which uses SisCone algorithm, SisCone algorithm is preferred over Anti-k_T [104]. There is no difference in the η and ϕ position of leading jet reconstructed by two algorithms, but the p_T of the jet reconstructed with SisCone algorithm is higher than that reconstructed with the Anti-k_T approach. These observations are clear from figure 4.8, which shows the correlation between leading jet reconstructed with SisCone and Anti-k_T algorithm. Figure 4.9 shows the comparison of the UE activity as a function of the leading track-jet clustered using SisCone and Anti-k_T algorithms. It is clear from the ratio plots that results are similar between the jet algorithms except at low p_T, where the p_T of the leading track-jet is different between the two algorithms.

Only charged tracks with $p_T > 0.5$ GeV and $|\eta| < 2.5$ are used for reconstructing track-jets. Leading track-jets are required to have $p_T > 1.0$ GeV and $|\eta| < 2.0$. Figure 4.10 shows the comparison of p_T and η of leading track-jet for data and MC predictions. There is a same level of agreement between data and MC as observed for 7 TeV analysis.

Figures 4.11–4.14 show the particle and Σp_T density, at detector level, as a function of leading track-jet p_T in the various transverse regions. The event selection criterion leaves 11M, 50K and 23K events in the ranges $1 \leq p_T < 25$, $25 \leq p_T < 50$ and $p_T \geq 50$ respectively of leading track-jet p_T. Qualitatively, variation in the UE activity with track-jet p_T is the same as that observed in other UE measurements. UE activity first shows a fast rise due to an increase in MPI contribution, and for track-jet p_T larger than ~ 8 GeV activity flattens because of MPI saturation. As with UE observations in other measurements, there is a visible slow increase with jet p_T in the transverse activity because of increasing radiation contribution. The activity in the transMAX region (Fig. 4.12) shows a faster rise in activity with leading track-jet p_T as compared to the transMIN (Fig. 4.13) and overall transverse region (Fig .4.11). This can be attributed to the transMAX region containing a third jet component or in general more radiation. This is clearly seen in the transDIF activity in figure 4.14, where the MPI activity has been subtracted from the transMAX region. The transMIN activity, being more sensitive to MPI shows an obvious flattening out into the plateau region, especially for the particle density.

Measurements are also compared with predictions by PYTHIA6 Z2 and PYTHIA8 4C Monte-Carlo. PYTHIA6 Z2 typically shows higher activity than both data and 4C, except for transDIF $\sum p_T$ density, where PYTHIA6 Z2 shows pretty good agreement up to about 50 GeV. The Z2 Monte Carlo overestimates data for the rest of the regions. PYTHIA8 4C generally underestimates data.

4.3.3 Data Correction

In order to compare UE data to various MC generated predictions and check for model and parameter differences, measured observables are fully corrected for detector effects and selection efficiencies to reflect the activity from all primary charged-particles in the same kinematic domain. Jets at generator level are clustered by applying the SisCone algorithm to charged-particles. Phase

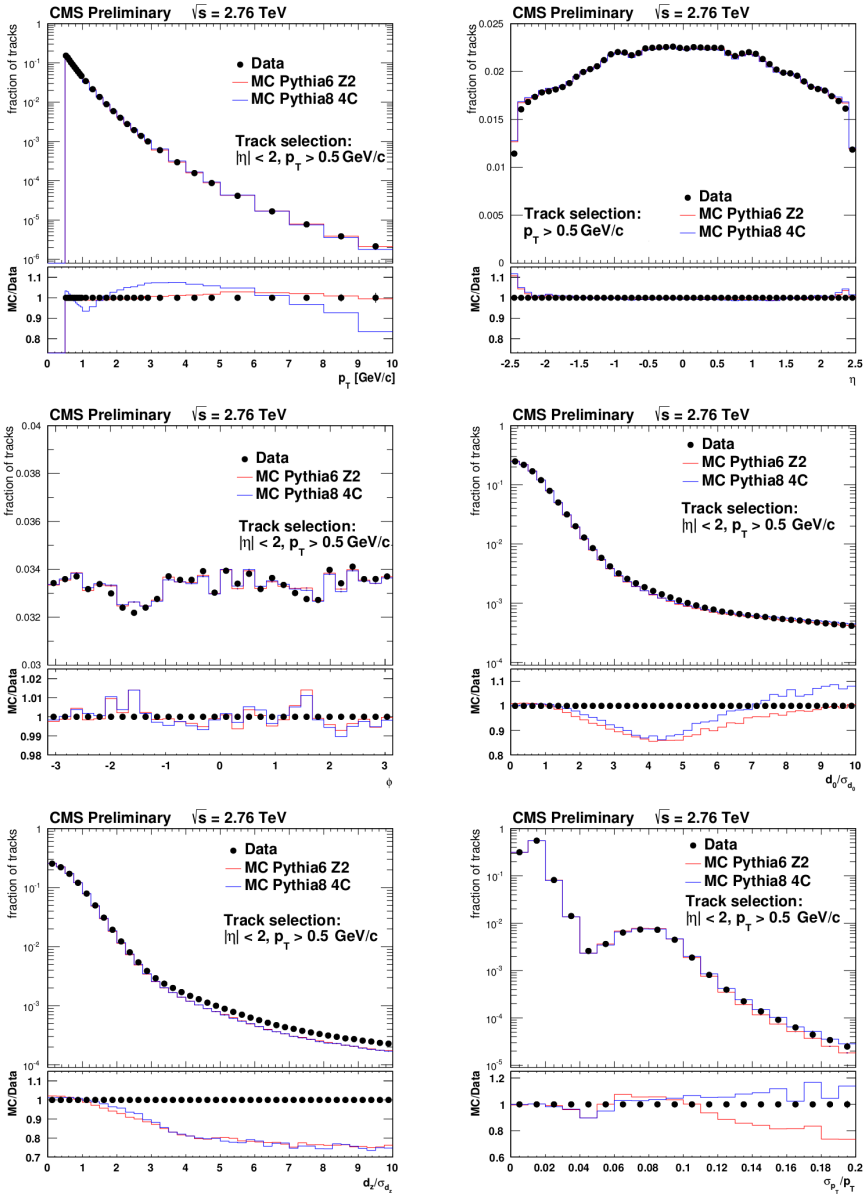

Figure 4.6 Comparison of track properties for data and Monte Carlo prediction PYTHIA6 Z2 and PYTHIA8 4C (top left) p_T, (top right) η, (centre left) ϕ, (centre right) d_0/σ_{d_0}, (bottom left) d_z/σ_{d_z}, (bottom right) σ_{pT}/p_T. Bottom panel in each plot shows the ratio of the MC prediction and the measurement to quantify the level of agreement.

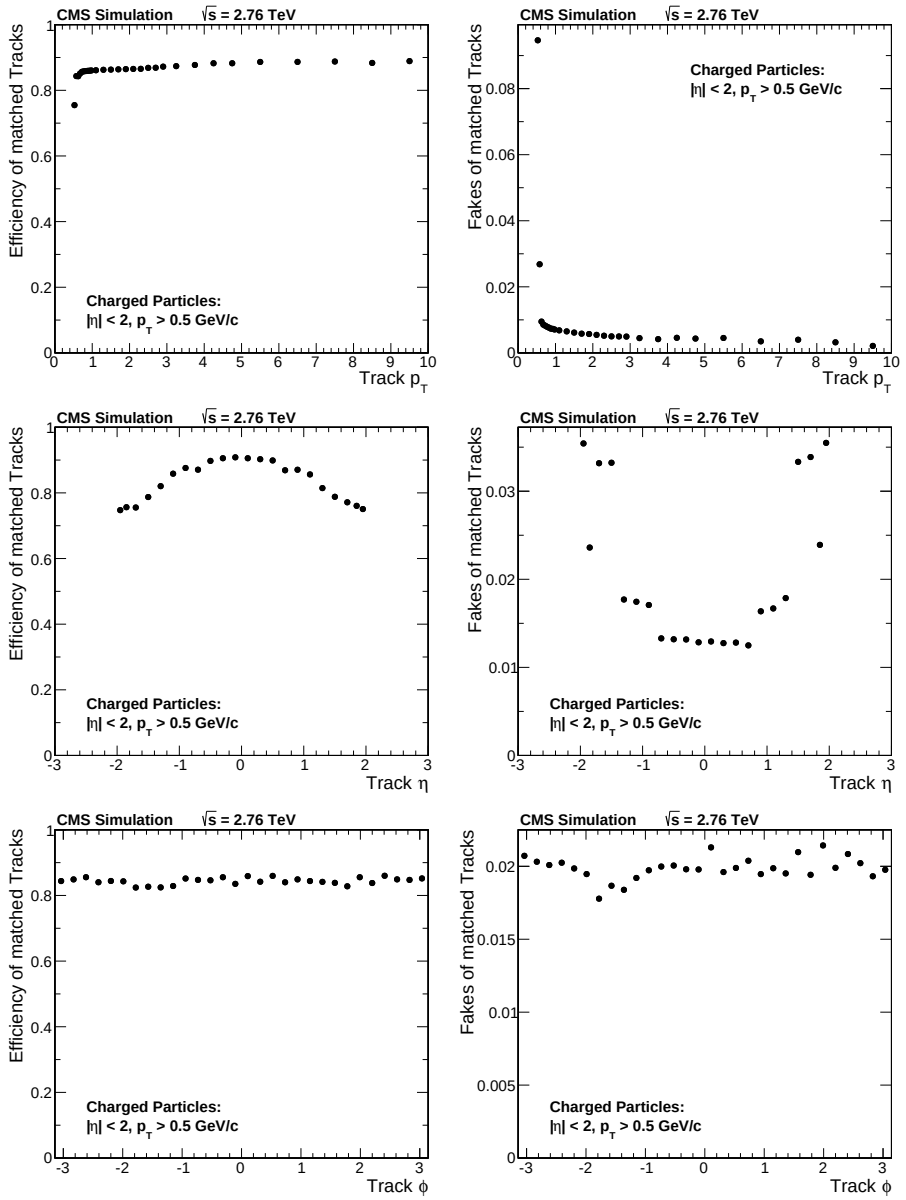

Figure 4.7 Efficiency (left column) and fake rate (right column) for reconstructed tracks as a function of track (top) p_T, (centre) η, and (bottom) ϕ obtained using PYTHIA6 MC.

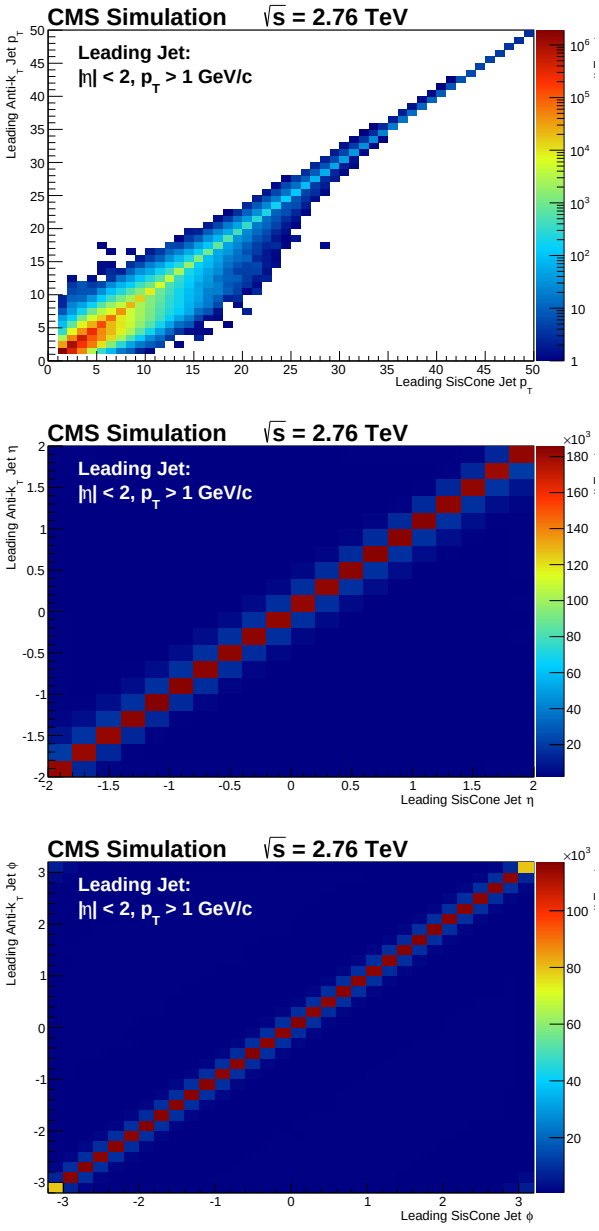

Figure 4.8 Correlation between leading jet reconstructed with the SisCone and Anti-k_T algorithms (top) jet p_T, (centre) jet η and (bottom) jet ϕ using PYTHIA6 MC.

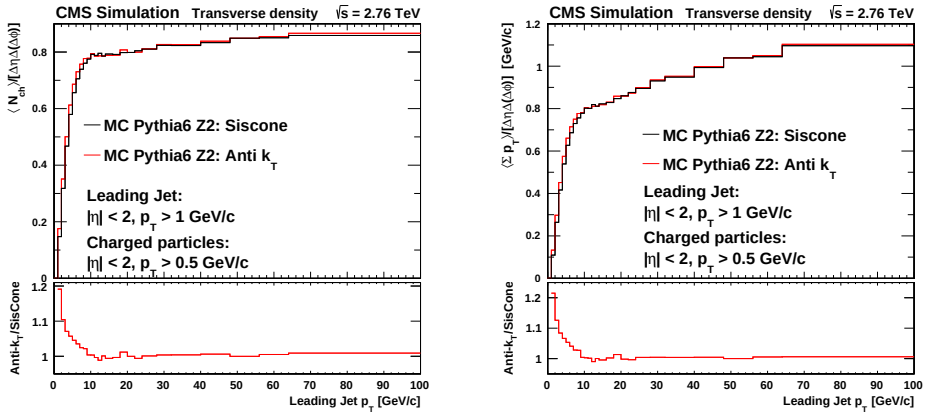

Figure 4.9 Comparison of UE activity as a function of leading track-jet clustered using SisCone and Anti-k_T algorithms for (left) transverse particle density and (right) transverse Σp_T density . Bottom panel in both plots shows the ratio UE activity as a function of leading track-jet clustered using SisCone and Anti-k_T algorithms.

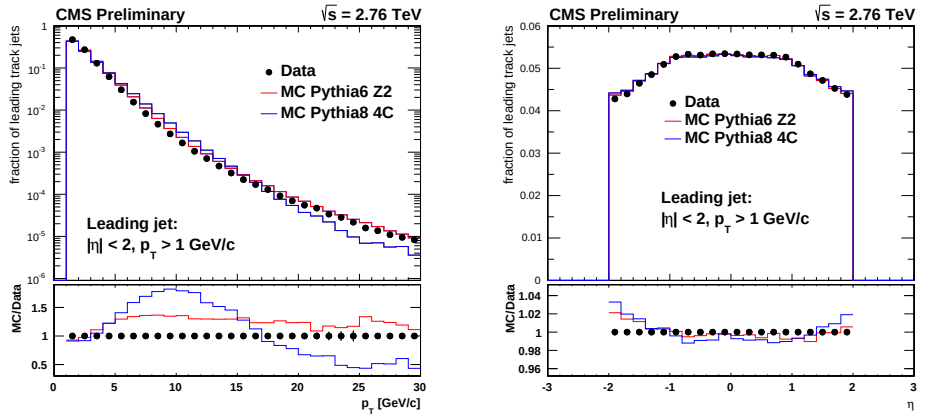

Figure 4.10 Comparison of track-jet properties for data and prediction PYTHIA6 Z2 (left) p_T, and (right) η. Bottom panel in each plot shows the ratio of the MC prediction and the measurement to quantify the level of agreement.

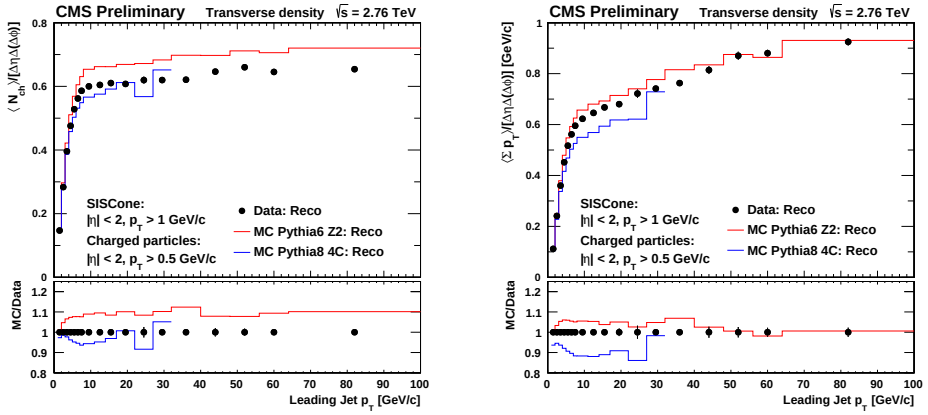

Figure 4.11 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) total transverse particle density and (right) total transverse Σp_T density as a function leading track-jet p_T.

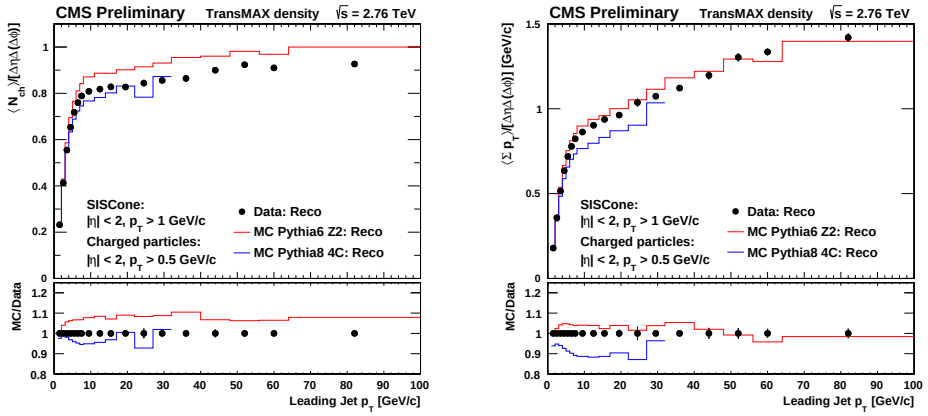

Figure 4.12 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transMAX particle density and (right) transMAX Σp_T density as a function leading track-jet p_T.

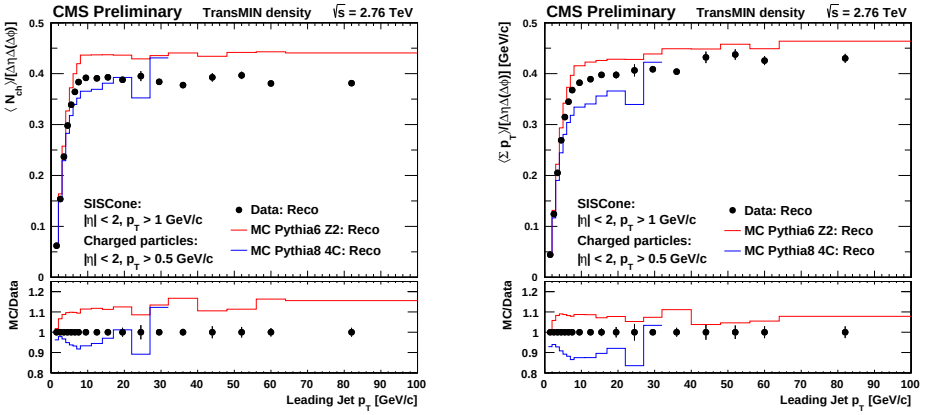

Figure 4.13 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transMIN particle density and (right) transMIN Σp_T density as a function leading track-jet p_T.

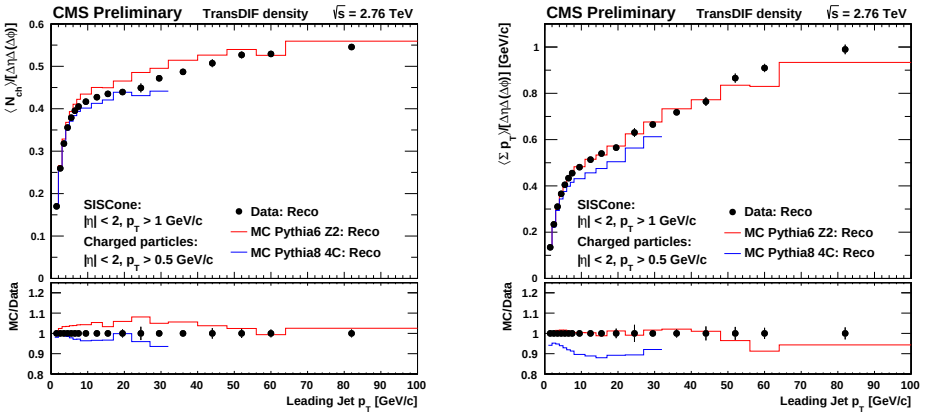

Figure 4.14 Comparison of UE activity for data, PYTHIA6 Z2 and PYTHIA8 4C at detector level; (left) transDIF particle density and (right) transDIF Σp_T density as a function leading track-jet p_T.

charged jet	$p_T > 1$ GeV and $\|\eta\| < 2.0$
charged particle	$p_T > 0.5$ GeV and $\|\eta\| < 2.0$

Table 4.3 Phase space definition for charged jet and charged particles at generator level.

space selection at particle level is summarised in table 4.3. Corrections are done with unfolding, employing the Bayesian method [105] inbuilt in the *RooUnfold* package. This requires the construction of a 4-Dimensional response matrix, characterising migrations between detector and generator level observables. The profile distributions are most conveniently expressed as 2-Dimensional histograms for unfolding, before averaging (profiling) along the axis representing the UE activity as a function of the leading jet p_T. Figures 4.15–4.16 shows the correlation between various kinematic quantities of leading jet and figures 4.17–4.20 show the correlation between track activity at detector and generator level. The leading jet matrices are almost diagonal because track-jets are well measured objects whereas track multiplicity and scalar Σp_T have large bin-to-bin migration. The Bayesian method properly corrects for these bin-to-bin migrations. The 4-Dimensional response matrix is constructed using simulated sample generated with PYTHIA6 Z2. Closure test is performed by splitting PYTHIA6 Z2 Monte-Carlo events into two halves, one for testing and one for training. There is a nice agreement, with a maximum deviation of 2–5% between unfolded and generator level distribution depending on the transverse region. This is shown in figures 4.21–4.24.

In the unfolding procedure the regularization parameter (number of iterations), is optimised by checking the variations in χ^2 with different number of iterations. Figure 4.25 shows the variation in χ^2 with different number of iterations for the particle and Σp_T density distributions. The regularization parameter for which variation in χ^2 is less than of 15% w.r.t. previous iteration is considered for the final results. As regularization depends on statistics and number of binning, the MC statistics and binning used are the same as that of data distributions. According to this requirement, 8 iterations are found to be optimal.

Unfolded results obtained with the Bayesian method are also compared with corrected results obtained by a simple bin-by-bin correction as shown in figure 4.26. The simple correction method, which does not account for bin migrations, consists of a 3 step procedure as follows:

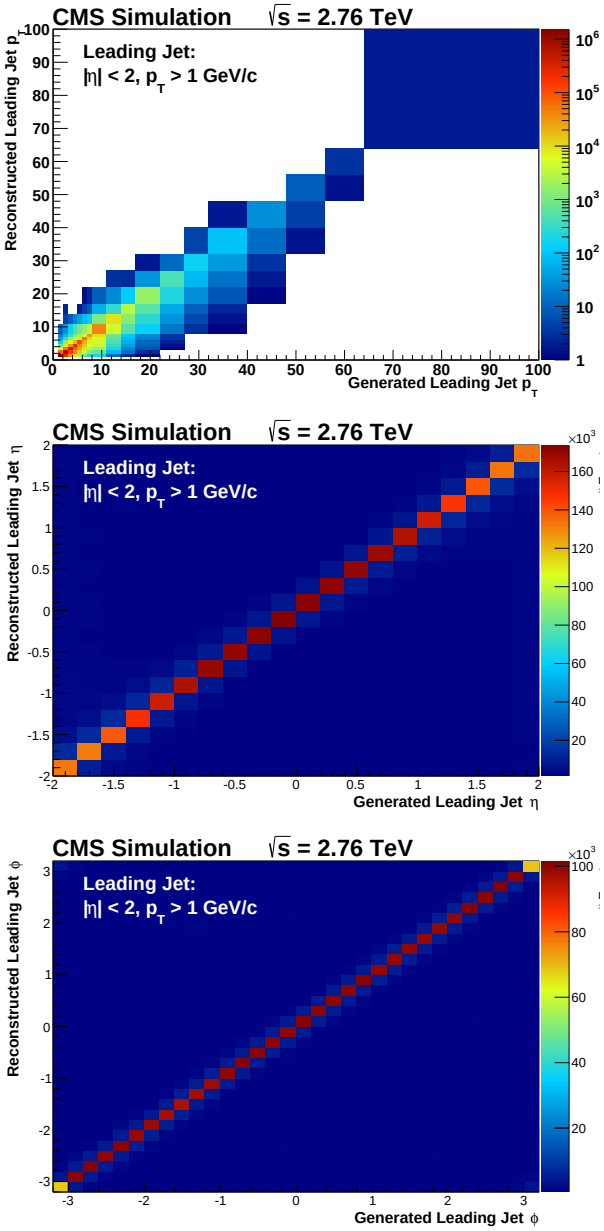

Figure 4.15 Correlation matrices for jet properties at reconstructed and generator level; (top) leading jet p_T, (centre) leading jet η and (bottom) leading jet ϕ. These are obtained using PYTHIA6 MC sample.

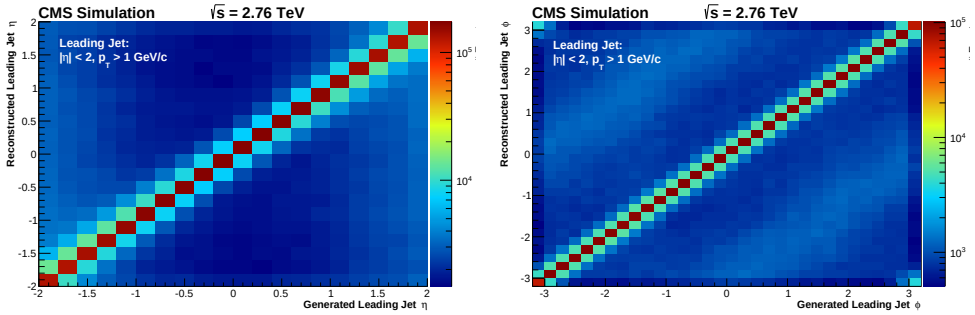

Figure 4.16 Correlation matrices for jet properties at reconstructed and generator level; (left) leading jet η and (right) leading jet ϕ with the z-axis in logarithmic scale. These are obtained using PYTHIA6 MC sample.

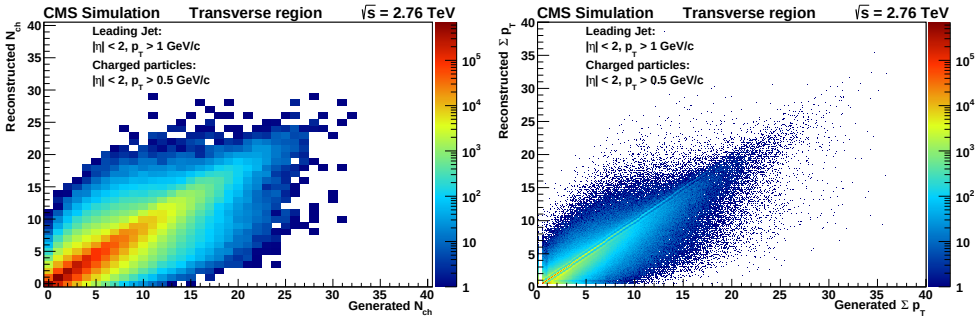

Figure 4.17 Correlation matrices for track properties at reconstructed and generator level; (left) total transverse track multiplicity and (right) total transverse Σp_T. These are obtained using PYTHIA6 MC sample.

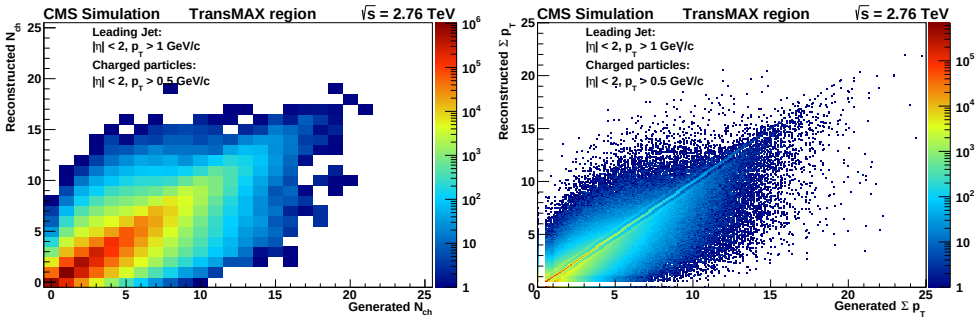

Figure 4.18 Correlation matrices for track properties at reconstructed and generator level; (left) transMAX track multiplicity and (right) transMAX Σp_T. These are obtained using PYTHIA6 MC sample.

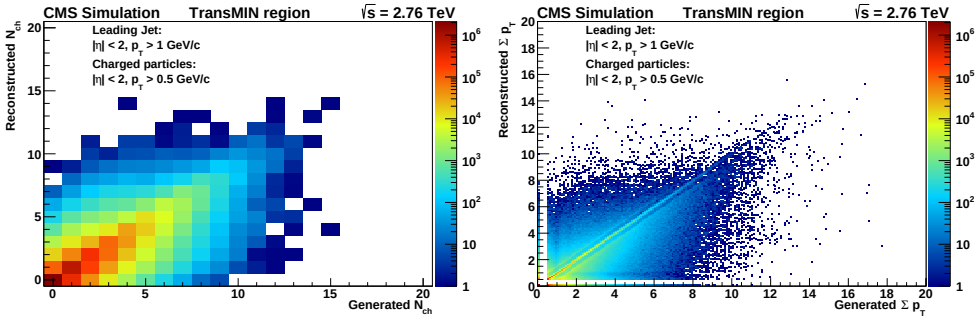

Figure 4.19 Correlation matrices for track properties at reconstructed and generator level; (left) transMIN track multiplicity and (right) transMIN Σp_T. These are obtained using PYTHIA6 MC sample.

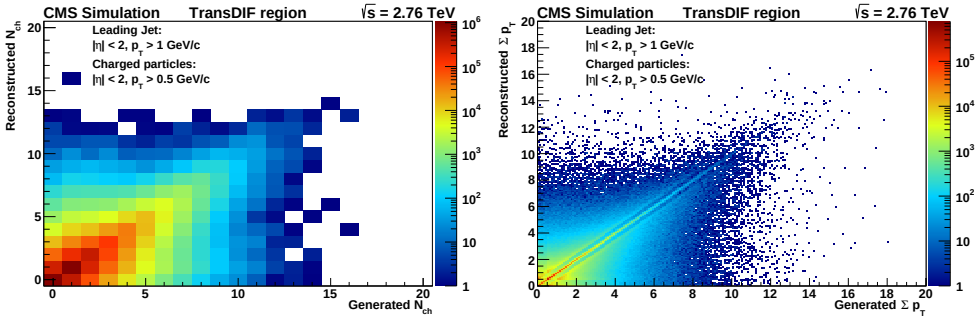

Figure 4.20 Correlation matrices for track properties at reconstructed and generator level; (left) transDIF track multiplicity and (right) transDIF Σp_T. These are obtained using PYTHIA6 MC sample.

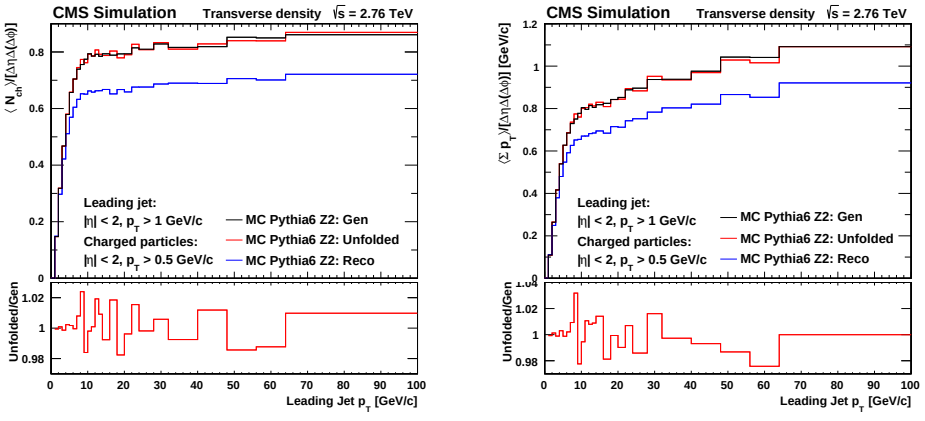

Figure 4.21 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transverse particle density and (right) transverse Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample.

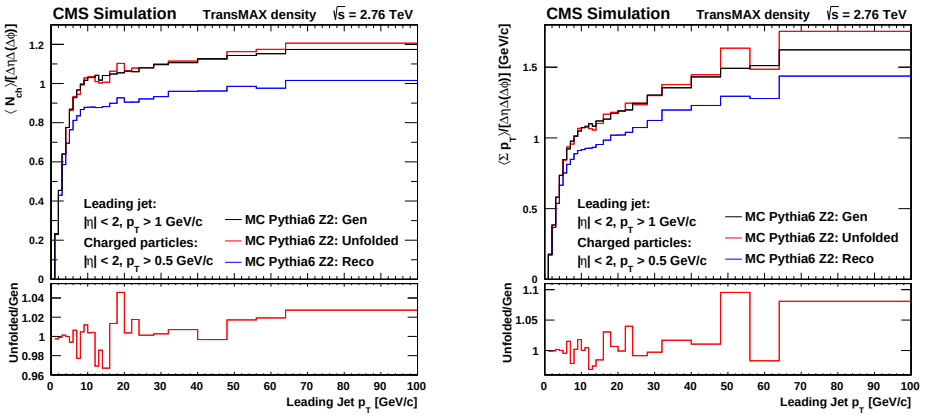

Figure 4.22 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transMAX particle density and (right) transMAX Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample.

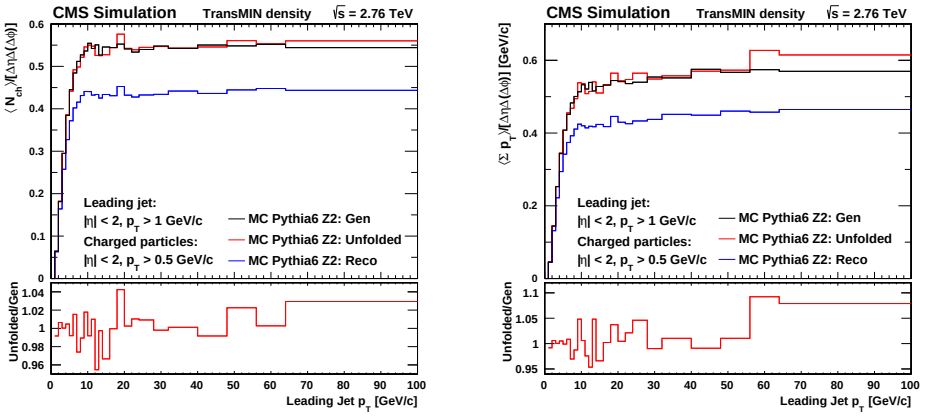

Figure 4.23 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transMIN particle density and (right) transMIN Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample.

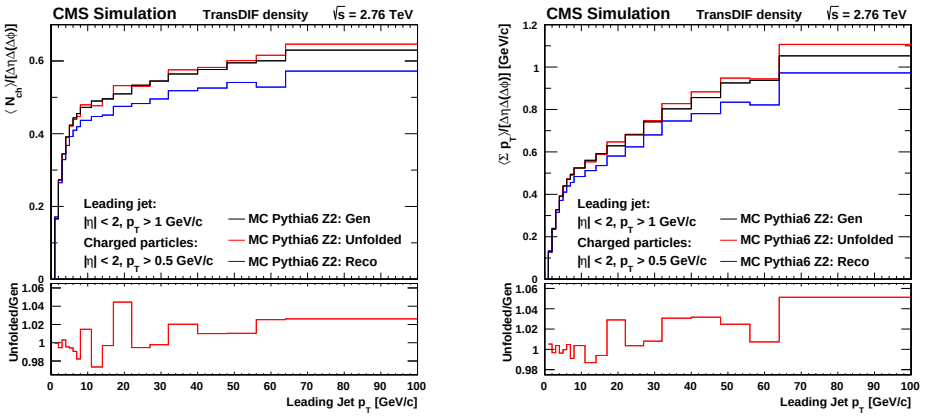

Figure 4.24 Closure test for unfolding; comparison of detector level, generator level and unfolded profile distributions of (left) transDIF particle density and (right) transDIF Σp_T density as a function of leading jet p_T. Bottom panel shows the ratio of unfolded and generator level profile distributions. These are obtained using PYTHIA6 MC sample.

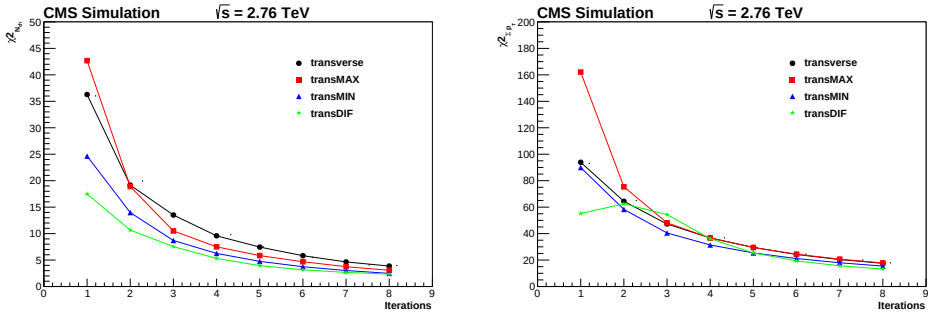

Figure 4.25 χ^2 of unfolding at different iterations for PYTHIA6 MC unfolded with an independent sample of PYTHIA6 MC; comparison is done for the various transverse regions of (left) particle density and (right) Σp_T density.

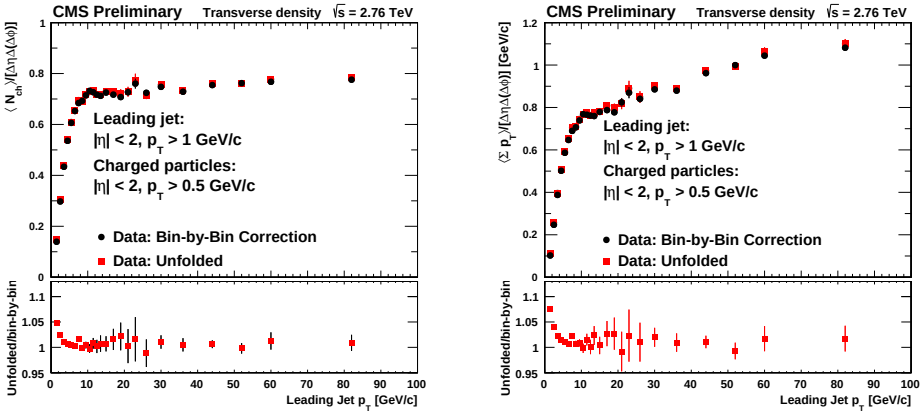

Figure 4.26 Comparison of unfolded results (left) particle density, (right) Σp_T density obtained by Bayesian method and simple bin-by-bin correction. Bottom plot shows the ratio of profile distributions corrected using Bayesian method and bin-by-bin correction.

1. Each track is weighted by the factor $\frac{1-fake(p_T,\eta)}{efficiency(p_T,\eta)}$, to obtain the corrected N_{ch} and $\sum p_T$. The track weighted profiles $\langle N_{ch} \rangle$ and $\langle \sum p_T \rangle$ is then calculated.

2. Bin-by-bin correction is applied to correct for mis-alignment or other effects contributing to a discrepancy between detector and generated level leading jets.

3. Finally, a last bin-by-bin correction is made for event selection.

Corrected results are very similar in the high jet p_T region, whereas for very low p_T jets (less than 5 GeV) there is deviation up to 5% due to large bin-migration involved in this region. Effects of the model dependency of the correction method are discussed in the next section, and are evaluated by using different MCs for the response matrix.

4.3.4 Systematic uncertainties

Various sources of systematic uncertainties are considered and discussed in this section. Systematic uncertainties vary bin-to-bin so quantitative numbers are quoted only for track-jet p_T equals to 20 GeV, where it is well into the plateau region.

- **Model dependence**: Systematic uncertainty due to model dependence is estimated by comparing unfolded results obtained using different MCs, i.e. PYTHIA6 Z2 for response matrix and PYTHIA8 4C for distributions to be unfolded. Figures 4.27–4.30 show the unfolded PYTHIA8 4C densities for the various regions, compared to the generated level 4C densities. Bottom panels show the ratio of unfolded profile distributions to the generated level distributions, which are taken as systematic uncertainties. At leading jet $p_T > 10$ GeV, event statistics start to diminish, contributing the huge fluctuations. Thus systematic uncertainties are taken bin-by-bin up to 10 GeV, whereas in higher p_T region, a constant fit is performed to avoid the fluctuations. This is shown by the blue dashed line in the bottom panels. Systematic uncertainty is estimated to about 1.2% for particle density and 1.6% for Σp_T density in the transverse plateau region. Values for the other regions are summarised in table 4.4.

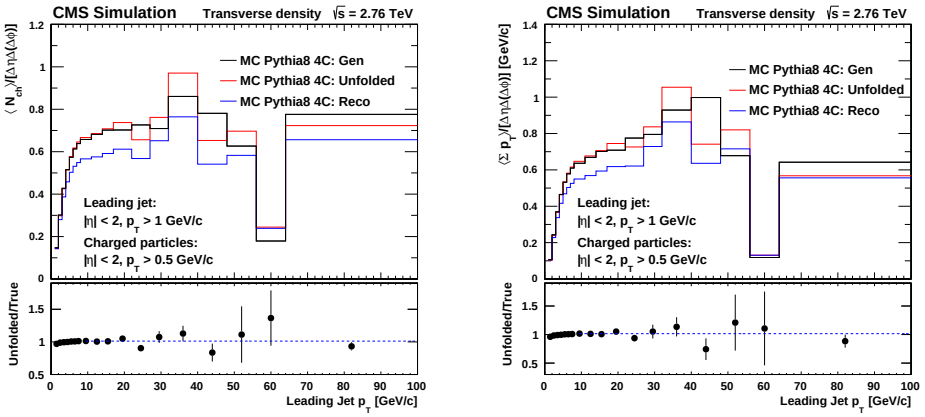

Figure 4.27 Unfolded PYTHIA8 4C profile distributions of (left) transverse particle density and (right) transverse Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.012 with $\chi^2/NDF = 14.26/10$ and (right) 1.016 with $\chi^2/NDF = 7.93/10$.

- **Trigger**: A systematic of 1.0% due to the minimum bias trigger has been estimated in [82], which adopts the same data set, selections and cuts but studies the UE activity in forward rapidities. Since the trigger systematic introduces uncertainties at the event level, it does not bias the topological definition of the UE activity. The systematic has been quantified for

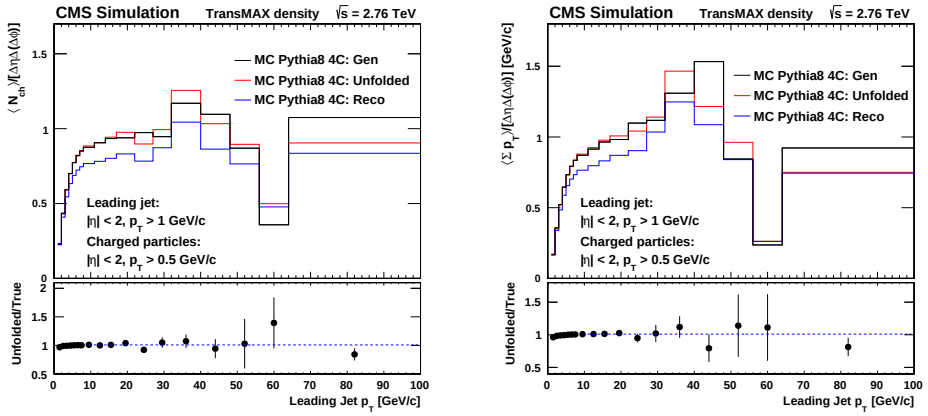

Figure 4.28 Unfolded PYTHIA8 4C profile distributions of (left) transMAX particle density and (right) transMAX Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.008 with $\chi^2/NDF = 9.58/10$ and (right) 1.009 with $\chi^2/NDF = 4.81/10$.

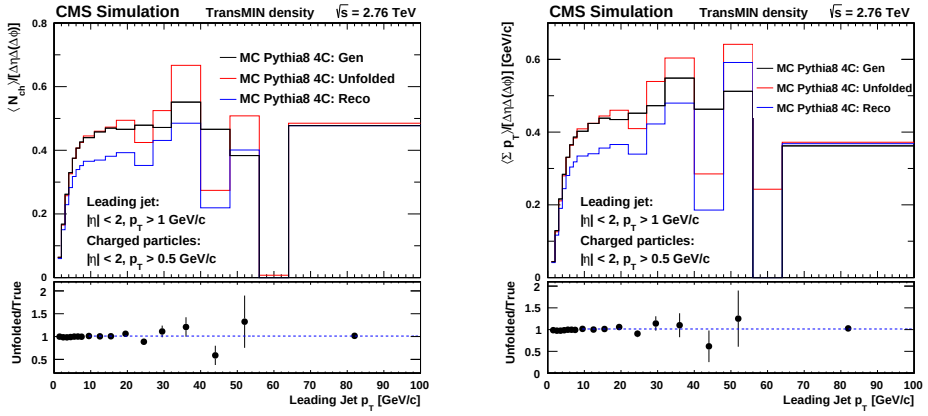

Figure 4.29 Unfolded PYTHIA8 4C profile distributions of (left) transMIN particle density and (right) transMIN Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 1.011 with $\chi^2/NDF = 11.41/9$ and (right) 1.014 with $\chi^2/NDF = 5.96/9$.

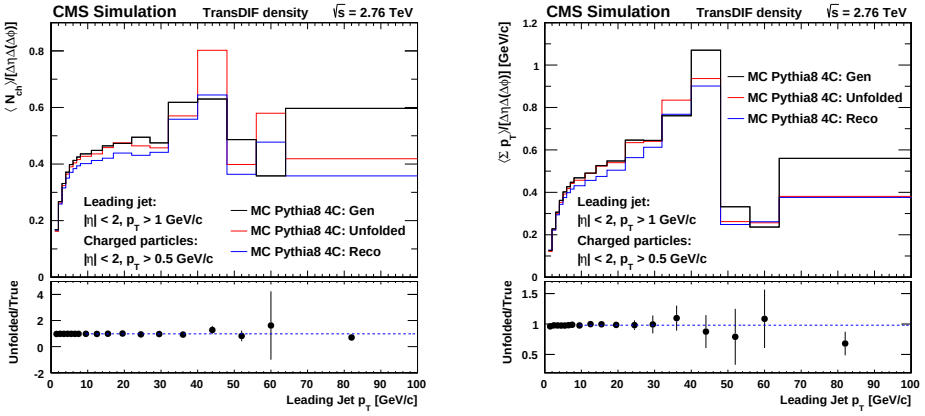

Figure 4.30 Unfolded PYTHIA8 4C profile distributions of (left) transDIF particle density and (right) transDIF Σp_T density, obtained using PYTHIA6 Z2 for the response matrix. Bottom panels represent the ratio of the Unfolded over generated level distributions. The blue dashed line is a best fit constant of (left) 0.979 with $\chi^2/NDF = 5.11/10$ and (right) 0.980 with $\chi^2/NDF = 7.27/10$.

minimum bias trigger and we have applied the uncertainty to Jet20 and Jet40 triggered samples as well.

Another effect of trigger efficiency uncertainties on the unfolded distributions is estimated. The uncertainty of the Jet20 efficiency w.r.t. the minimum bias trigger is $\sim 5\%$ (Fig. 4.5(bottom left)) and the uncertainty of the Jet40 trigger w.r.t the Jet20 trigger is $\sim 1\%$. By varying the event weights of the Jet20 sample by $\pm 5\%$ and performing unfolding, the uncertainty can be evaluated by comparing with the original distributions. This results in an uncertainty of less than 0.3%. Figures 4.31–4.32 compares the distributions with weighted Jet20 sample with un-weighted ones.

- **Impact parameter significance**: Different cuts on d_z and d_0 significance are studied by considering unfolded profiles and distributions. The ratios of the corrected profiles are directly computed considering two different cuts on impact parameter significance: less than 2 and 3, where the latter is the standard analysis cut. This variation in selection criteria affects profile distributions by 2.8–3.0%.

- **Fake tracks contribution**: The number of fake tracks, originating from secondaries and combinatorics, are different between data and simulated samples. To study the possible effect of this difference, the amount of

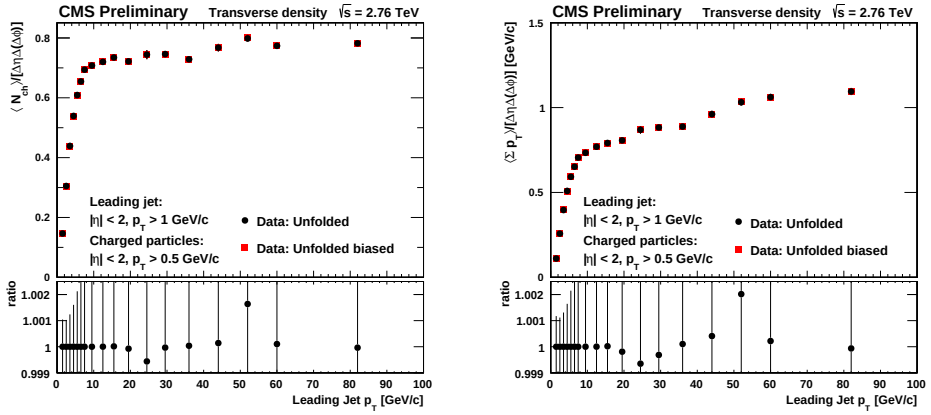

Figure 4.31 Unfolded data profile distributions of (left) transverse particle density and (right) transverse Σp_T density where Jet20 events are weighted by 1.05 and compared to the normal distribution. Bottom panels represent the ratio of the biased over un–biased distributions.

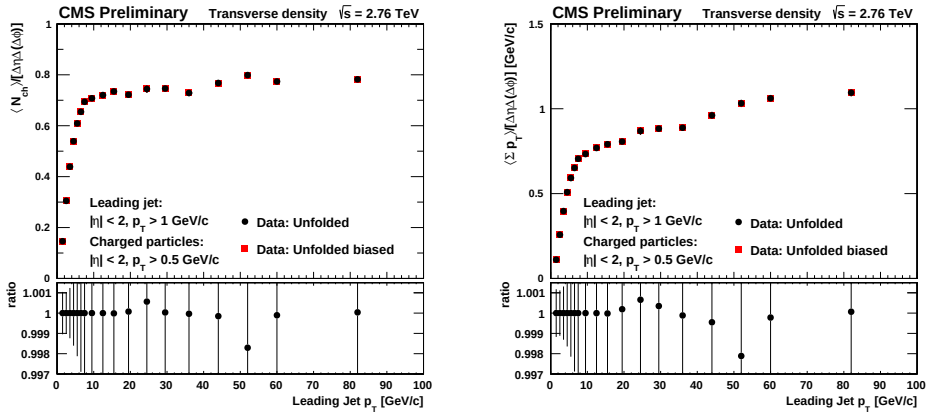

Figure 4.32 Unfolded data profile distributions of (left) transverse particle density and (right) transverse Σp_T density where Jet20 events are weighted by 0.95 and compared to the normal distribution. Bottom panels represent the ratio of the biased over un–biased distributions.

fake is varied by 50% up and down. This variation affects the profile distributions by 0.4–0.5%.

- **Vertex selection**: In order to evaluate the bias induced by the vertex cut, based on dof, the analysis observables are re-computed in 2 different scenarios changing the request on the number of degrees of freedom to 4 ± 1. The profile distributions are affected by 0.2% with this variation in vertex selection criteria.

- **Track selection**: The effect of variation in track selection is also considered as one of the systematic. We define a reference selection, where High-Purity tracks are selected with an impact parameter significance smaller than 3 sigma and $p_T/\sigma_{p_T} < 5\%$. In the alternative selection, the High-Purity requirement is replaced by a simpler condition relying only on the number of crossed layers: $N_{layers} \geq 4$ and $N_{pixellayers} \geq 2$, respectively for silicon and pixel detectors. This variation affects the profile distributions by 0.2–0.3%.

- **Pileup**: The effect of pileups on the measured activity is investigated following a method described in [106]. Average pileup in data can be estimated with the formula:

$$\langle pileup \rangle = \sum_{n=1}^{\infty} (n-1)\,\mathscr{P}(n) = \frac{\mu}{1 - exp(-\mu)} - 1, \qquad (4.3)$$

where the distribution of vertices $\mathscr{P}(n)$ is Poisson. By fitting a Poisson distribution, $\mu \approx 0.121$ and $\langle pileup \rangle \approx 0.062$, as shown in Fig 4.33 (left).

The fraction of merged and lost vertices when the pileup vertices cannot be resolved can be estimated by fitting a Gaussian distribution to the Δz distribution of 2 vertices shown in Fig 4.33 (right). The missing area at small Δz when comparing the Gaussian fit to the Δz distribution can be attributed to 2 vertices that have been merged into 1. The fraction of those vertices can be calculated by $f_{merged} = \frac{\delta z}{\sigma_z \sqrt{\pi}}$, where $\delta z \approx 0.4$ cm is the estimated value of Δz at which the vertices start to merge, and $\sigma_z \sqrt{2}$ is the standard deviation of the Gaussian fit. The value $f_{merged} \approx 0.043$.

The pileup systematic is estimated to be $\langle pileup \rangle \times f_{merged} \approx 0.003$, multiplied again by the fractional residual. The fractional residual is calculated

by performing the same analysis of UE activity for events with 2 vertices, and taking the difference of this activity with that of 1 vertex events, divided by activity from 1 vertex events. This is found to be around 0.014% for transverse particle density and 0.018% for transverse Σp_T density averaged across leading track-jet p_T. The effect is an insignificant uncertainty on our results.

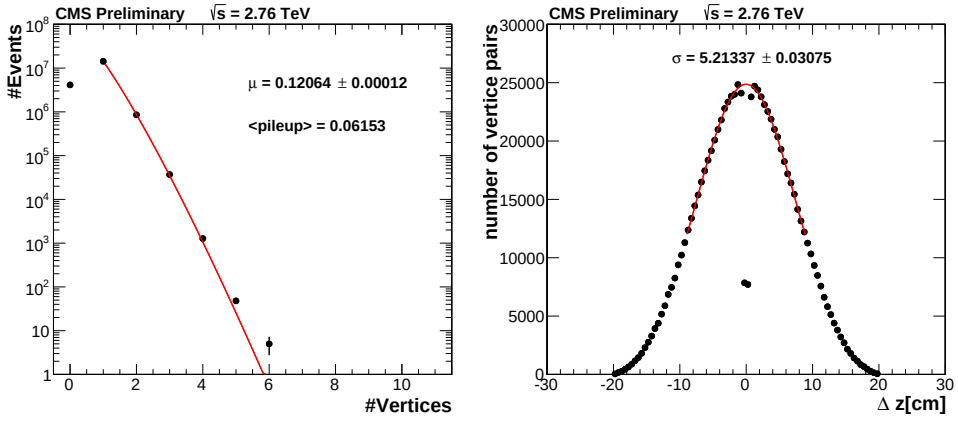

Figure 4.33 Estimation of the pileup effect with (left) the distribution of number of vertices with a Poisson fit and (right) the Δz distribution of 2 vertices.

Systematic effects related to detector conditions are taken from studies performed for the UE measurement at $\sqrt{s} = 7$ TeV, where conditions are the same as that for 2.76 TeV. These are briefly summarised below:

- **Material budget**: The relevant characteristic of the tracker material affecting the reconstruction performance is the thickness crossed by a particle, either in units of radiation length X_0, or in units of interaction length λ_i. The uncertainty related to the tracker material has been studied by comparing different scenarios: maximally (X_{0Max} and λ_{iMax}) and minimally affected (X_{0Min} and λ_{iMin}). This is obtained by making variations in the material densities, resulting in an uncertainty which affects the profile distributions by 1%.

- **Dead channels**: Since data from different runs are combined in this analysis, the effect of possible run by run variations in the dead channel maps has been studied comparing with a simulation with no dead channels at all. A conversion factor of 0.05 is applied to be realistic with the measured

fraction of dead channels, corresponding to a 5% variation of the dead channel map during the data taking. This variation in the number of dead channels affects the profile distributions by 0.1%.

- **Tracker alignment**: Tracker alignment impacts track reconstruction performance, which is critical to the analysis. The effects of geometrical misalignment of the pixel and silicon detectors have been studied by comparing the alignment expected at start-up conditions to the one that should be achieved by the design performance. Tracker mis-alignment is found to affect profile distributions by 0.2–0.3%.

- **Beamspot**: The beamspot is not perfectly centered in the detector. In particular, significant z offsets can reduce the acceptance of the tracker system. This effect has been studied by simulating different beam spot schemes compatible with the variation over the analyzed runs. The resulting contribution of the systematic uncertainty on the profile distributions is measured to be 0.2%.

Table 4.4 summarises the various systematic uncertainties (in percentage) for the profile distributions at $p_T^{jet} = 20$ GeV. Except for model dependence, all systematics have been adopted to be the the same across the various transverse regions. Trigger, impact parameter significance, vertex selection and pileup are all systematics coming from event level fluctuations. These have no topological dependence and will have the same systematic uncertainty regardless of the topological region. Fake rates and track selection takes care of individual track performance and also has no region specific preference. Finally, material budget, dead channels, tracker alignment and beamspot are systematics coming from detector conditions which have no region specific bias. Thus only model dependence will have a varied systematic uncertainty across the different regions. Figures 4.34–4.37 show the total and individual systematic biases for the profile distributions.

Source	*transverse*		*transMAX*		*transMIN*		*transDIF*	
	$\langle N_{ch} \rangle$	$\langle \sum p_T \rangle$	$\langle N_{ch} \rangle$	$\langle \sum p_T \rangle$	$\langle N_{ch} \rangle$	$\langle \sum p_T \rangle$	$\langle N_{ch} \rangle$	$\langle \sum p_T \rangle$
Model dependence	1.2	1.6	0.8	0.9	1.1	1.4	2.1	2.0
Trigger	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Impact parameter sig.	2.8	3.0	2.8	3.0	2.8	3.0	2.8	3.0
Fake	0.5	0.4	0.5	0.4	0.5	0.4	0.5	0.4
Vertex selection	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Track selection	0.3	0.2	0.3	0.2	0.3	0.2	0.3	0.2
Pileup	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Material budget	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Dead channels	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Tracker alignment	0.2	0.3	0.2	0.3	0.2	0.3	0.2	0.3
Beamspot	0.2	0.2	0.2	0.2	0.2	0.2	0.2	0.2
Syst. Total	3.4	3.7	3.3	3.5	3.4	3.8	3.9	3.9

Table 4.4 Summary of the systematic uncertainties (in percentage) due to various sources and in the various transverse regions.

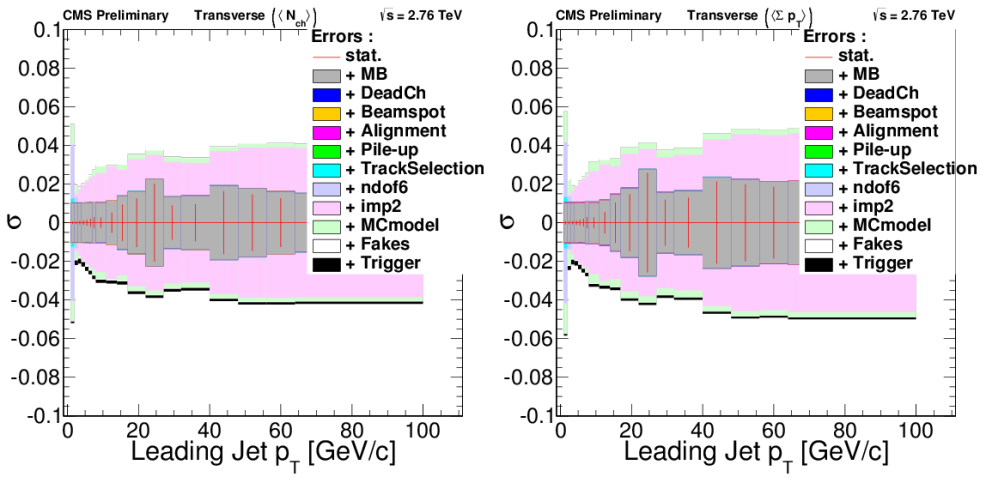

Figure 4.34 Summary of the systematic uncertainties for the profile distributions panel (left) for transverse particle and (right) for transverse $\sum p_T$ densities.

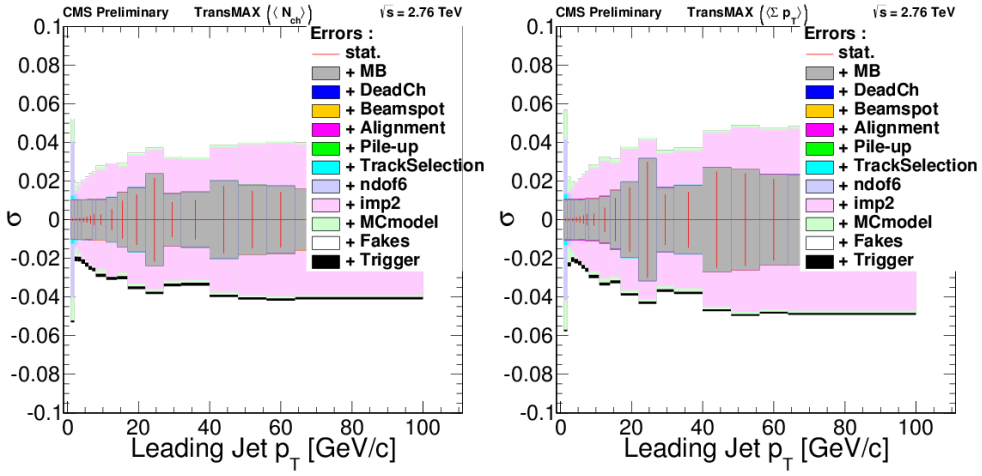

Figure 4.35 Summary of the systematic uncertainties for the profile distributions panel (left) for transMAX particle and (right) for transMAX $\sum p_T$ densities.

4.3.5 Results

This section presents and discusses the key distributions from this analysis, the particle density $\langle N_{ch} \rangle / [\Delta\eta\Delta(\Delta\phi)]$ and Σp_T density $\langle \sum p_T \rangle / [\Delta\eta\Delta(\Delta\phi)]$ as a function of p_T^{jet}. For these distributions full unfolding was performed, correcting for detector effects and selection efficiencies.

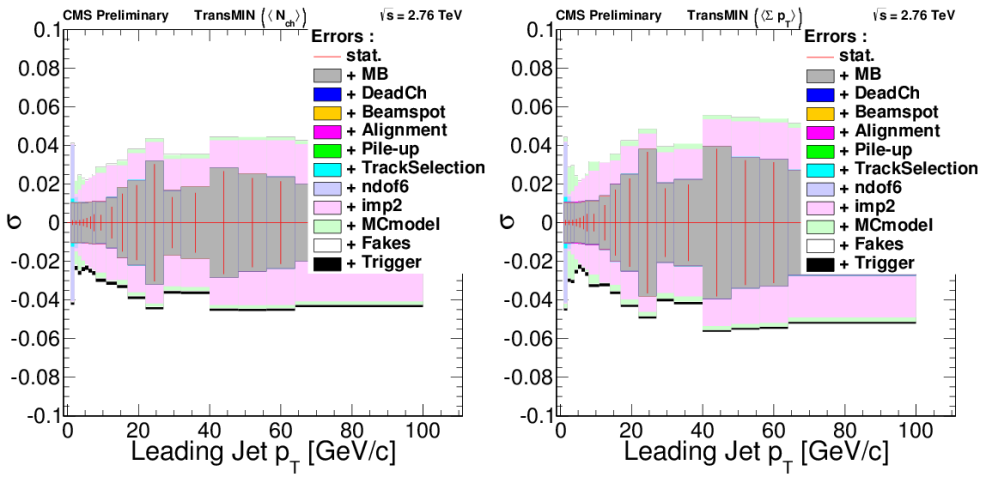

Figure 4.36 Summary of the systematic uncertainties for the profile distributions panel (left) for transMIN particle and (right) for transMIN $\sum p_T$ densities.

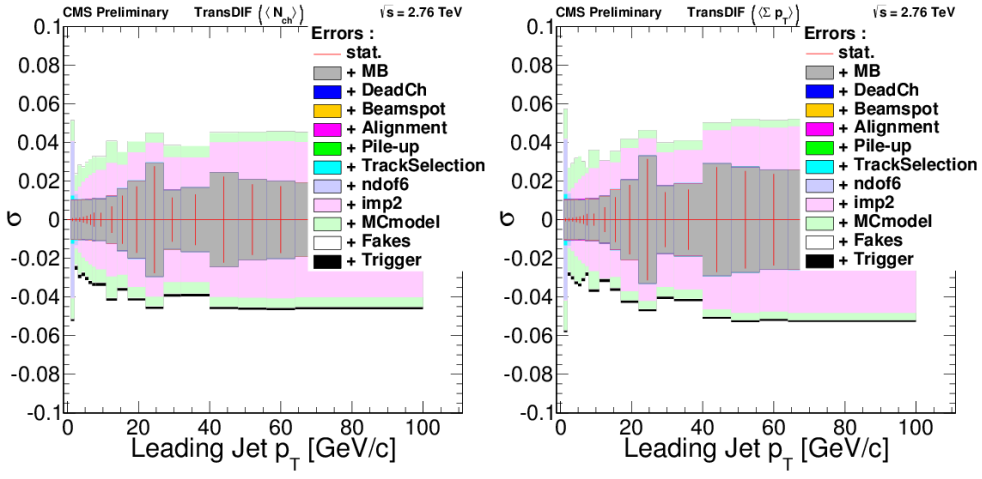

Figure 4.37 Summary of the systematic uncertainties for the profile distributions panel (left) for transDIF particle and (right) for transDIF $\sum p_T$ densities.

In figure 4.38, the particle (top) and Σp_T (bottom) densities in the transverse region as a function of p_T^{jet} are shown. Both measured distributions rise steeply for $p_T^{jet} \lesssim 8$ GeV, attributed mainly to the increase of MPI activity, followed by a plateau-like region with nearly constant particle density and a slow increase of $\sum p_T$ density, indicating an increase of the amount of energy delivered per particle as the scale of the interaction increases. The data are compared with various MC predictions. In general, all PYTHIA6 and PYTHIA8 tunes predict the distinctive change of the amount of activity as a function of p_T^{jet}, except for PYTHIA8 4C which undershoots the data for $p_T^{jet} > 5$ GeV. The latest PYTHIA6 (PYTHIA8) tune CUETP6S1 (CUETP8S1) improves the description of data in comparison with the previous tune Z2* (4C). In general, all PYTHIA6 and PYTHIA8 tunes describe the data to within 5–10%. The HERWIG++ UE-EE-5C tune also provides a fair description of the data.

In figures 4.39 and 4.40, the transMAX and transMIN particle (top) and Σp_T (bottom) densities as a function p_T^{jet} are shown. The transMIN activity is roughly half that of the transMAX activity. The shape is also quite different in both categories, at high-p_T the distributions show a slow rise in transMAX, while in transMIN the plateau-like region is much flatter. In the different MC predictions, the same tendency as in figure 4.38 is observed. The description of the data is generally slightly worse for transMIN than transMAX densities.

The transDIF densities are presented in figure 4.41. As discussed earlier, this distribution is expected to be sensitive to the final- and initial-state radiations. Both particle (Fig. 4.41 (top)) and Σp_T densities (Fig. 4.41 (bottom)) show a rise with p_T^{jet} and the plateau-like region at about $p_T^{jet} \simeq 8$ GeV as seen in the previous plots actually shows an obvious increase with p_T^{jet}. The various MC tunes seem to provide a better description of the data as compared to the transverse, transMAX, and transMIN densities.

From figures 4.39, 4.40, and 4.41 one can conclude that the activity from MPI and beam-beam remnant is almost independent on the hard scale for $p_T^{jet} \gtrsim 8$ GeV, while activity from initial- and final-state radiations increases with p_T^{jet}.

The center-of-mass energy dependence of the activity in the transverse region is presented in figure 4.42 as a function of p_T^{jet} for $\sqrt{s} = 0.9$, 2.76, and 7 TeV. A strong growth with increasing center-of-mass energy of the activity in the transverse region is observed for the same value of p_T^{jet}. All tunes predict the center-of-mass energy dependence of the activity well.

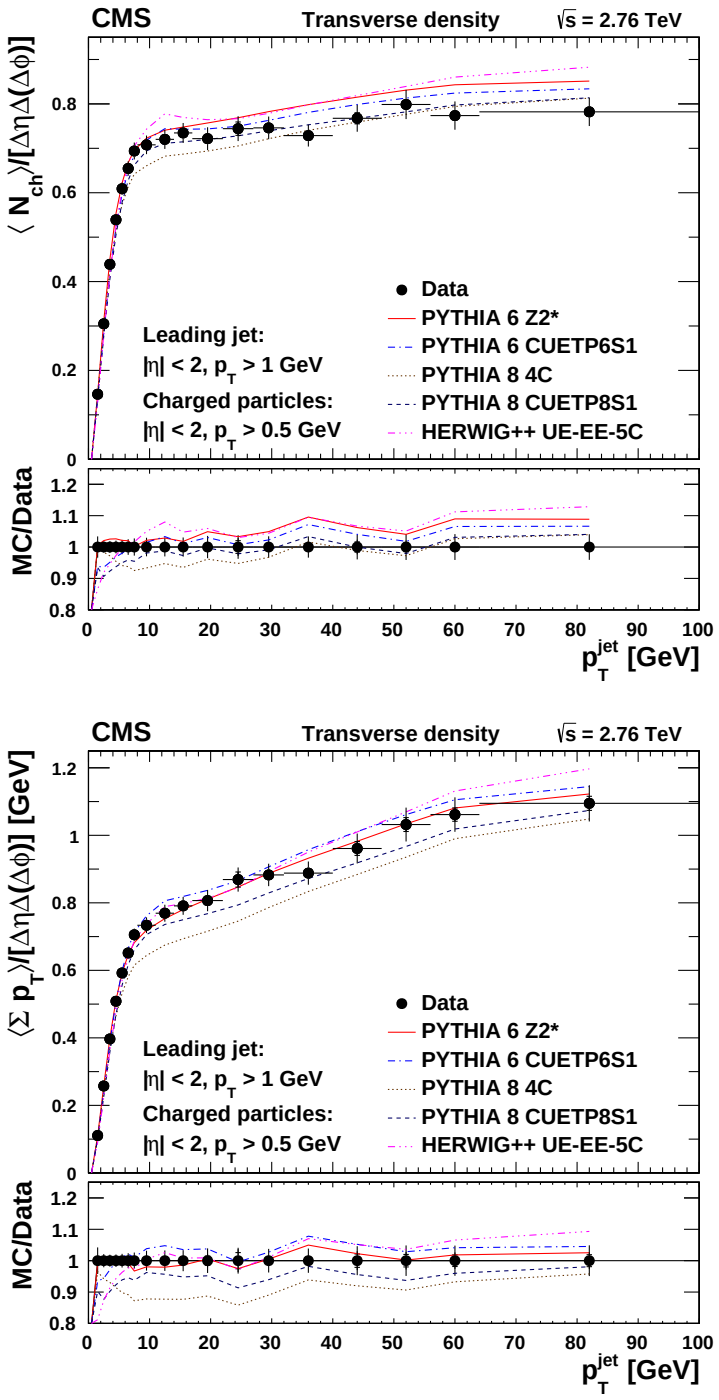

Figure 4.38 Measured (top) particle density, and (bottom) Σp_T density, in the transverse region relative to the leading charged-particle jet in the event ($|\eta| < 2, 60° < |\Delta\phi| < 120°$), as a function p_T^{jet}. The data (symbols) are compared to various MC simulations (curves). The ratios of MC simulations to the measurements are shown in the bottom panels. The inner error bars correspond to the statistical uncertainties, and the outer error bars represent the statistical and systematic uncertainties added in quadrature.

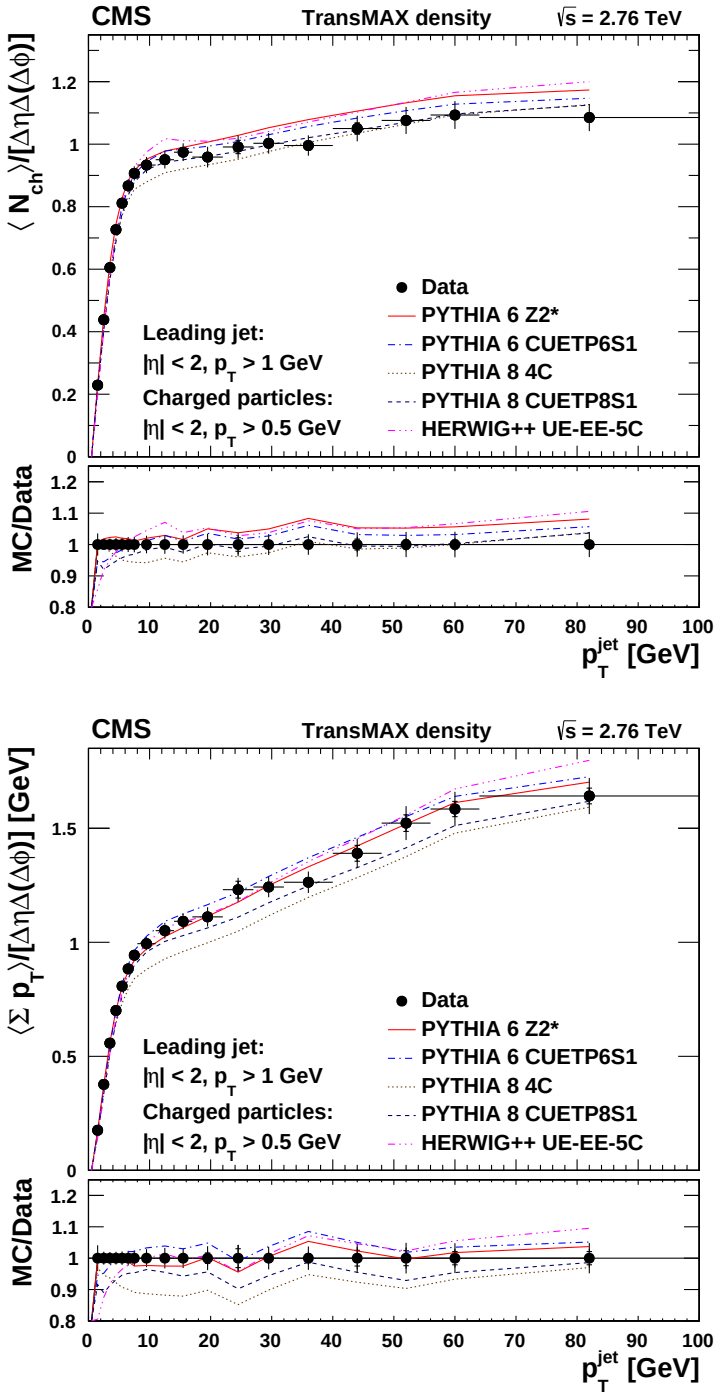

Figure 4.39 Measured (top) particle density, and (bottom) Σp_T density, in the transMAX region ($60° < |\Delta\phi| < 120°$, relative to the leading charged-particle jet in the event, with maximum UE activity), as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38.

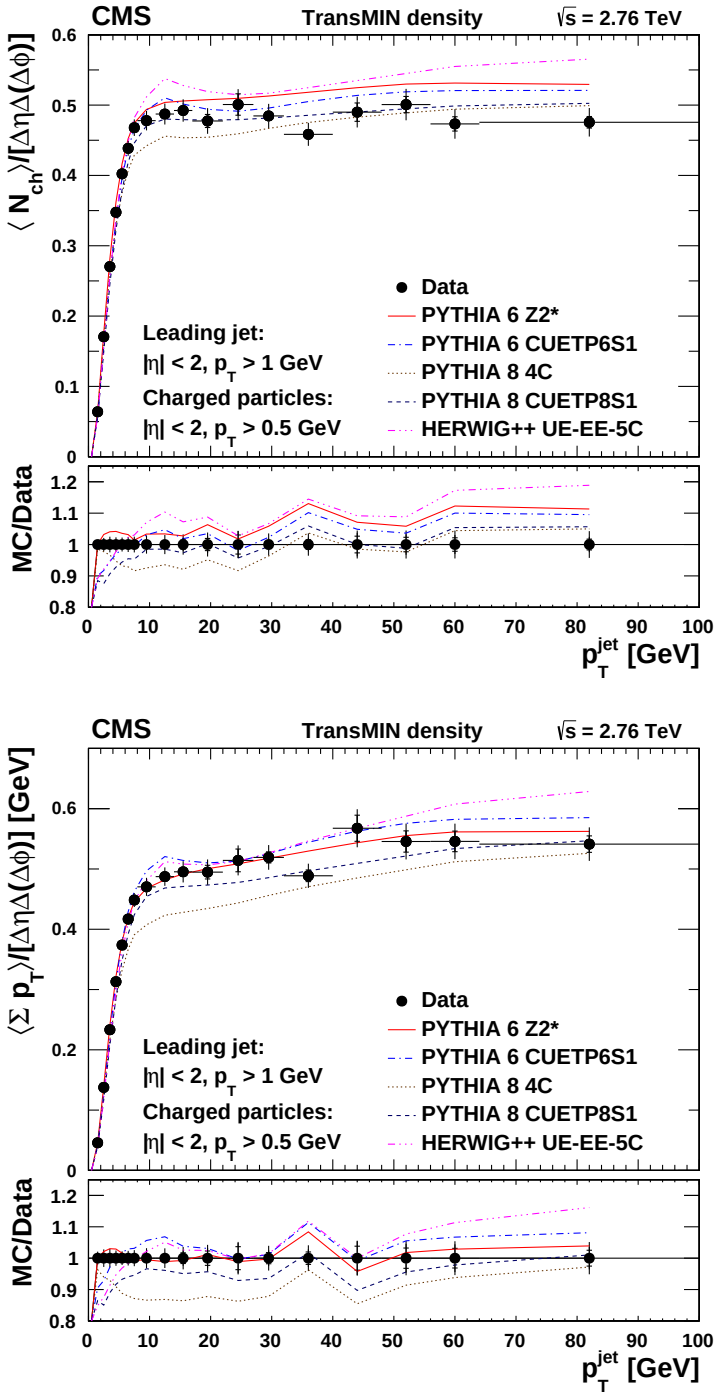

Figure 4.40 Measured (top) particle density, and (bottom) Σp_T density, in the transMIN region ($60° < |\Delta\phi| < 120°$, relative to the leading charged-particle jet in the event, with minimum UE activity), as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38.

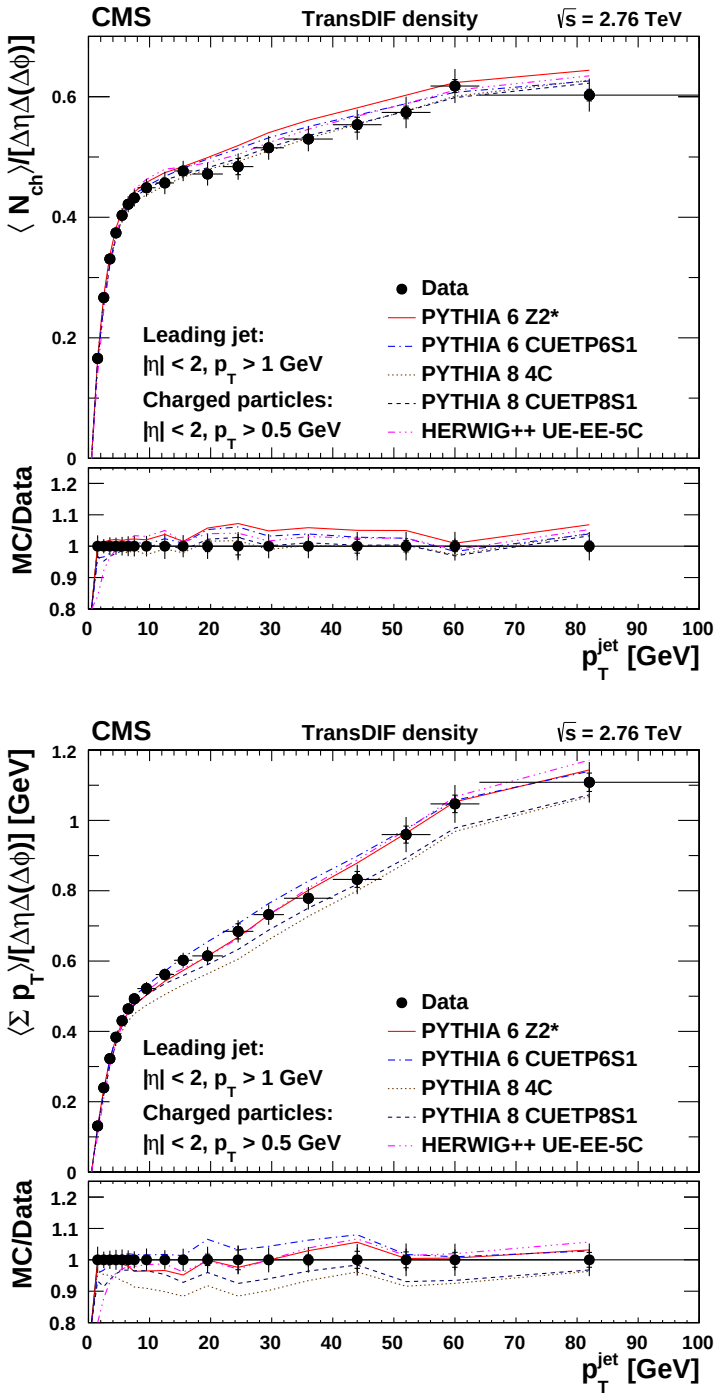

Figure 4.41 Measured transDIF activity (see text for its definition) for (top) particle density, and (bottom) Σp_T density, as a function of p_T^{jet}. The definitions of the symbols and error bars are the same as for Fig. 4.38.

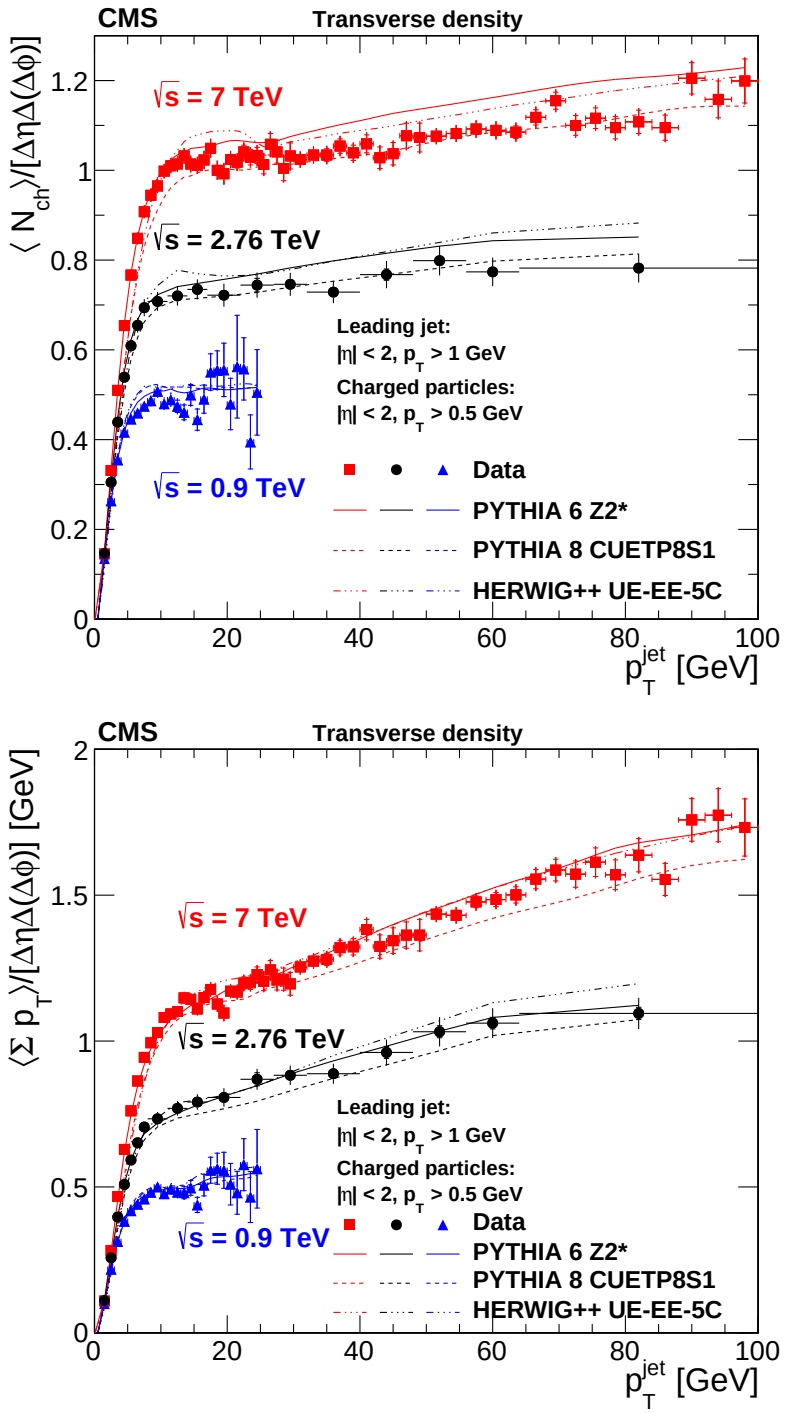

Figure 4.42 Comparison of UE activity at $\sqrt{s} = 0.9$, 2.76, and 7 TeVfor (top) particle density, and (bottom) Σp_T density, as a function of p_T^{jet} [14, 15]. The data (symbols) are compared to various MC simulations (curves). The definition of the error bars is the same as for Fig. 4.38.

4.4 Measurement of the UE at 13 TeV

A measurement of the UE activity is performed at $\sqrt{s} = 13$ TeV [68, 69] in the same way as the UE analysis at $\sqrt{s} = 2.76$ TeV. On top of a measurement of the UE activity performed using leading charged-particle jets as reference objects, leading charged-particles were also used. While the leading charged-particles and charged-particle jets are required to be produced in the central pseudorapidity region ($|\eta| < 2$), the charged-particle jets are selected with $p_T^{jet} \geq 1$ GeV and the charged-particle with $p_T \geq 0.5$ GeV.

The data sample forms part of the data, with low PU, collected by the CMS detector in July 2015. The collected data sample has an average of 1.3 simultaneous pp collisions per event and correspond to an integrated luminosity of 281 nb^{-1}. Events are required to pass the zerobias trigger, which only needs a coincidence of both BPTX devices, in contrast to the UE analysis at $\sqrt{s} =$ 2.76 TeV which also requires the BSC. This ensures that beam bunches have crossed the IP but not necessarily having pp collisions. Even so, events having pp collisions are enhanced as the PU condition is significantly higher than that for the UE analysis at 2.76 TeV. Undesirable effects from simultaneous interactions can be mostly removed by the requirement of a single vertex for selected events. The relatively higher PU also means that this event selection criterion will still leave sufficient statistics for the UE study with leading charged-particle (charged-particle jet) up to $p_T = 24$ ($p_T^{jet} = 50$) GeV.

In this analysis the MC generators PYTHIA8 [84] CUETP8M1 and Monash [107] tunes, HERWIG++ [85] CUETHS1 tune, as well as the EPOS 1.99 [108] are used. Events are simulated with no PU as well as with an average PU of 1.3, which is required for the data correction and calculation of the systematic uncertainties. The MC generator EPOS uses a similar string fragmentation hadronisation model [109] as that of Lund string hadronisation model [86] as used by PYTHIA. However, EPOS describes soft-parton dynamics by Gribov-Regge theory [109], with the exchange of virtual quasi-states as multi-pomeron exchanges. The hard-parton processes at high energies are described similarly but generalised to include hard-pomeron scattering, which are equivalent to a leading-order pQCD Dokshitzer-Gribov-Lipatov-Altarelli-Parisi (DGLAP) evolution approach as used by both PYTHIA and HERWIG++.

Track quality and object selection follow the same as that of the UE analysis at 2.76 TeV. The measured UE densities are then fully corrected for detector effects and selection efficiencies. Systematic uncertainties are quantified by a systematic variation in the different conditions that could give rise to a bias in the final results. In general, these vary between the different densities (transAVE/transDIF/transMAX/transMIN), and as a function of p_T (p_T^{jet}). The model dependency of the correction method gives an uncertainty of up to 4% (4%) above 5 GeV as a function of p_T (p_T^{jet}). In the lowest bin, deviation goes up to 8%. The systematic due to PU goes up to 2% (4%) depending on p_T (p_T^{jet}). Fake mis-modelling contributes to a systematic of up to 4% (2%) in the lowest p_T (p_T^{jet}) bin depending on the density type, and decreases to 1% above 2 GeV. The choice of impact parameter significance and vertex degree of freedom both contribute negligibly to the uncertainty.

The corrected transAVE (transverse), transMAX, transMIN, and transDIF UE densities are compared to previous UE measurements at $\sqrt{s} = 0.9$, 2.76, and 7 TeV as well as predictions of several MC event generators, providing constraints on the modelling of the UE dynamics.

4.4.1 Results

The corrected distributions of the average particle and Σp_T densities as a function of p_T/p_T^{jet} are compared with the predictions by several simulations. Figures 4.43 and 4.44 report the densities as a function of p_T^{jet}. The measured distributions for average Σp_T and particle densities are best described by the Monash tune of the PYTHIA8. Predictions by other simulations describe the measurements with a maximum deviation of up to 30%.

Figure 4.45 shows the average particle density, in the transverse region, as a function of the p_T of the leading charged-particle. The level of agreement between the measurements and the simulations is quantified by the ratio plots shown in the bottom panels. The measurements are better described by the Monash tune of PYTHIA8, while PYTHIA8 CUETP8M1 describe the measurements within 10–20%. Predictions by the CUETHS1 tune of HERWIG++ fails in the low p_T region. EPOS describes the low-p_T rising region well but fails to describe the plateau region, with a variation between 20–40%. Figure 4.46 reports similar distributions but for the average Σp_T density as a function of the

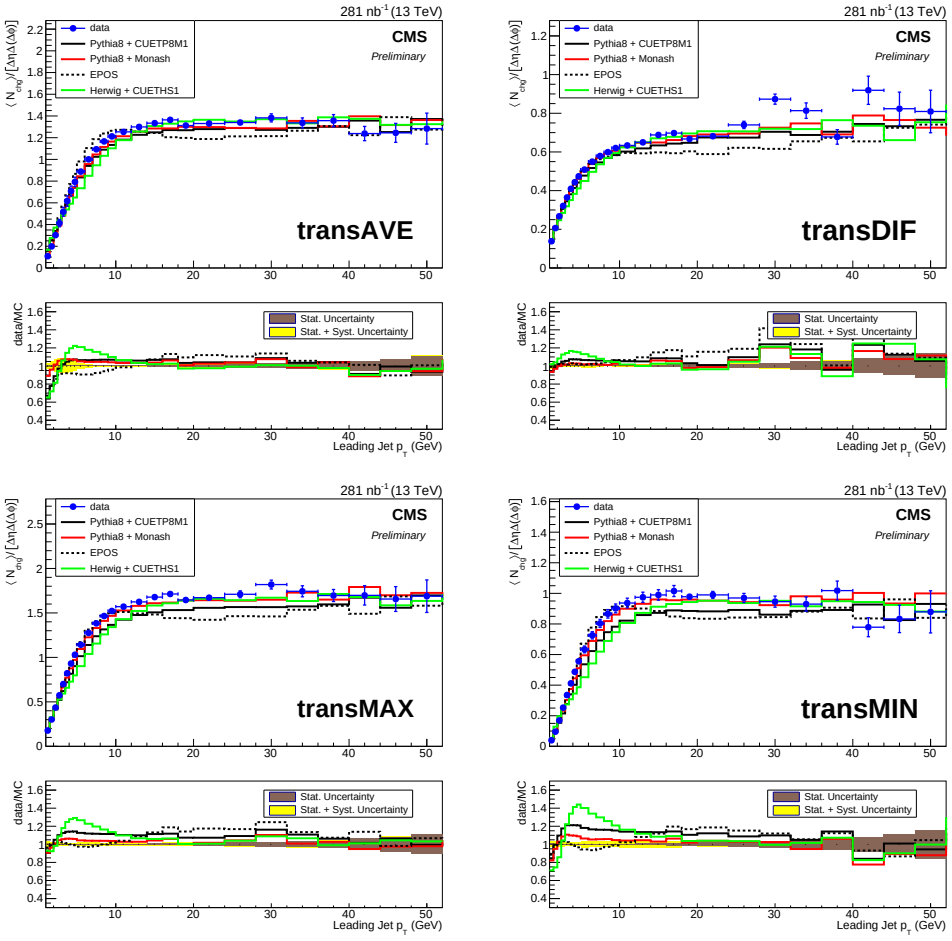

Figure 4.43 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average particle densities with the various simulations as a function of p_T^{jet}. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.

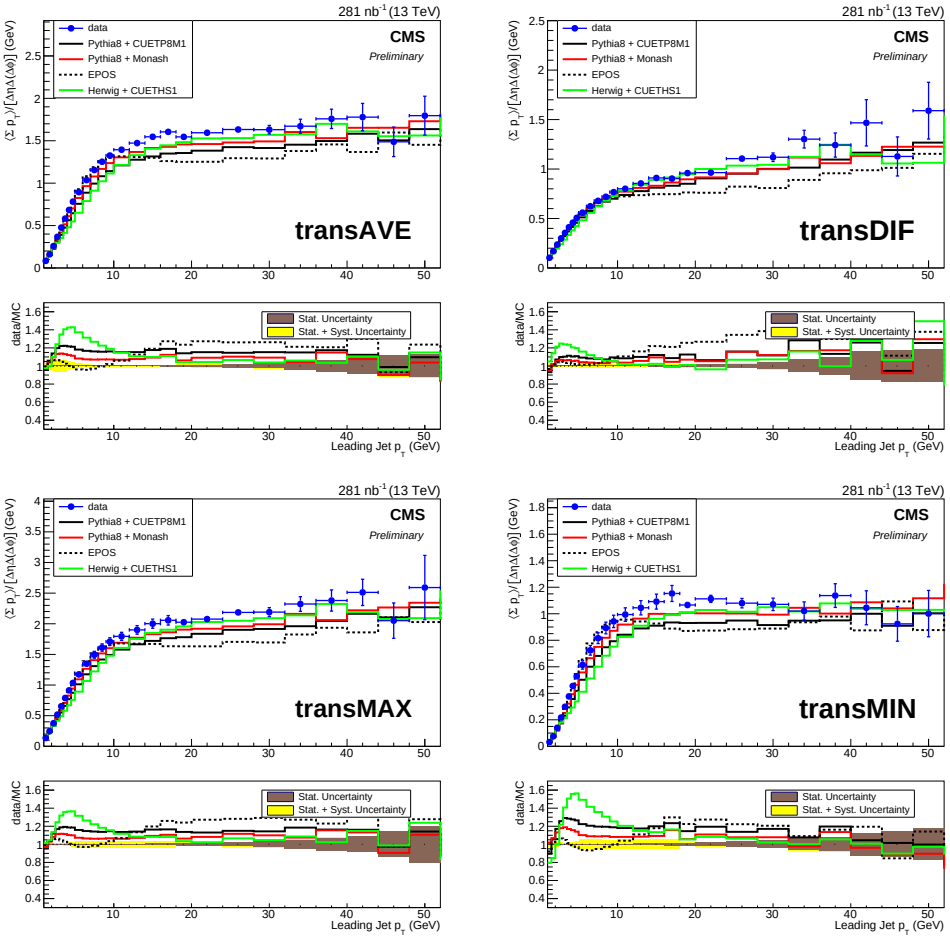

Figure 4.44 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average energy densities with the various simulations as a function of p_T^{jet}. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.

leading charged particle p_T. The amount of agreement of the measured average Σp_T density with simulations is the same as that for the average particle density, and the best description is again given by the Monash tune of PYTHIA8.

In all plots, the densities increase sharply with p_T (p_T^{jet}) up to 5 (12–15) GeV and rise slowly beyond. It is obvious from the evolution that the transMIN densities are flatter in the plateau region compared to the transMAX and transDIF densities which continue to show a slow rise. Simulations describe the same qualitative behavior of the measurements, i.e. the sharp rise followed by a flattening of the UE activity, and a larger rise in transMAX and transDIF in the plateau region compared to transMIN. The level of agreement between simulations and measurements fall within 10–20% in the plateau region but differ in the low p_T region. The sharp rise with p_T is interpreted in the MC models as due to an increase in the MPI contribution which saturates at high p_T. A persistent slow increase in the large p_T region for transMAX and transDIF densities is mainly due to an increase in the initial- and final- state radiation with p_T. As MPI activity is expected to be uniform in the whole phase-space, the transMIN densities capture mainly the activity coming from MPI whereas transDIF densities give the evolution of the radiation with p_T of the reference object.

Figures 4.47 and 4.48 show comparisons between various MC simulated samples and data across centre-of-mass energies of 0.9, 2.76, 7, and 13 TeV for transAVE densities, and at 2.76 and 13 TeV for transDIF, transMAX, and transMIN densities as a function of p_T^{jet}. There is a strong rise in the UE activity as a function of the centre-of-mass energy as predicted by the MC tunes. This is attributed to an increase in the of number partons with smaller fractional momenta x_L. The transMIN densities show a stronger $\sqrt{s}$ dependence than the transDIF density, indicating that the activity coming from MPI depends more strongly on $\sqrt{s}$ compared to ISR and FSR.

4.5 Summary

Measurements of the underlying event are performed for pp collisions at $\sqrt{s} =$ 2.76 and 13 TeV with the CMS experiment at the LHC. The analyses use events with charged-particles (for $\sqrt{s} = 13$ TeV) or charged-particle jets produced in the pseudorapidity region ($|\eta| < 2$) and with transverse momentum $p_T \geq 0.5$ GeV ($p_T^{jet} \geq 1$ GeV) for the leading charge-particle (charged-particle jet).

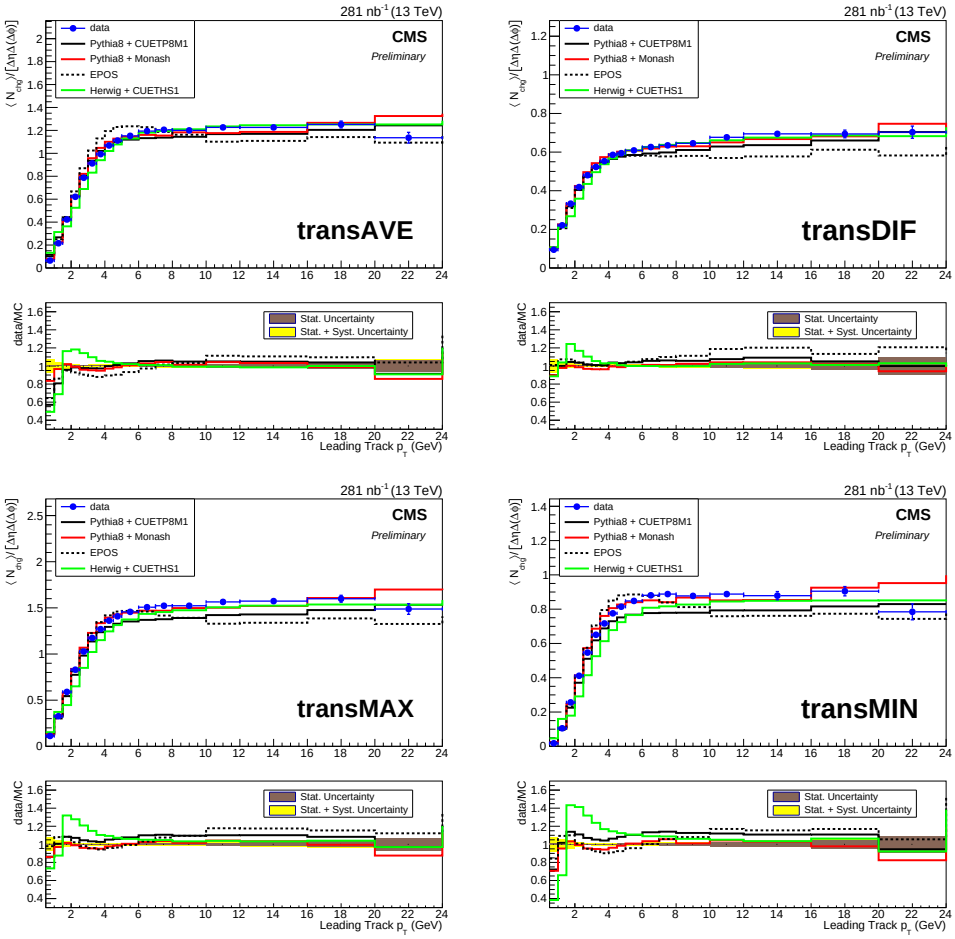

Figure 4.45 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average particle densities with the various simulations as a function of p_T. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.

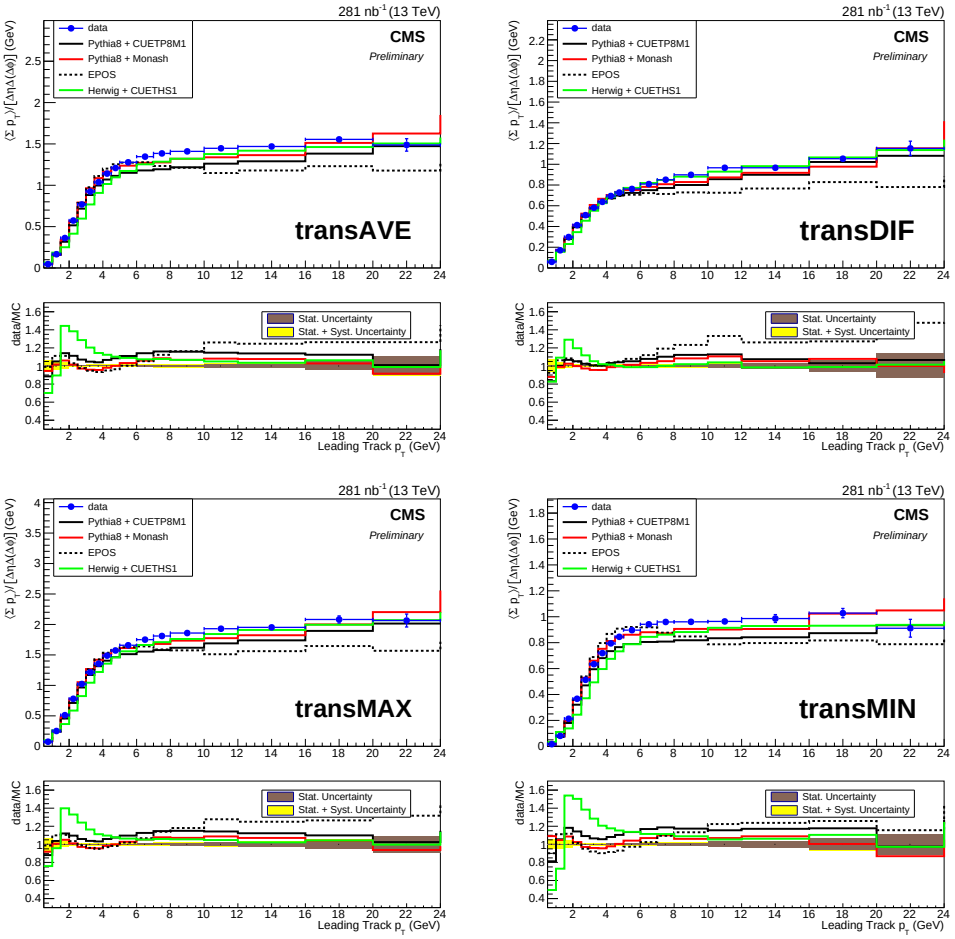

Figure 4.46 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN average energy densities with the various simulations as a function of p_T. The error bar represents the statistical and systematic uncertainties added in quadrature. Bottom panels shows the ratio of the simulation with the measurements. The brown band in the bottom plot represents the statistical uncertainty in the corrected data whereas the total uncertainty is shown in the yellow band.

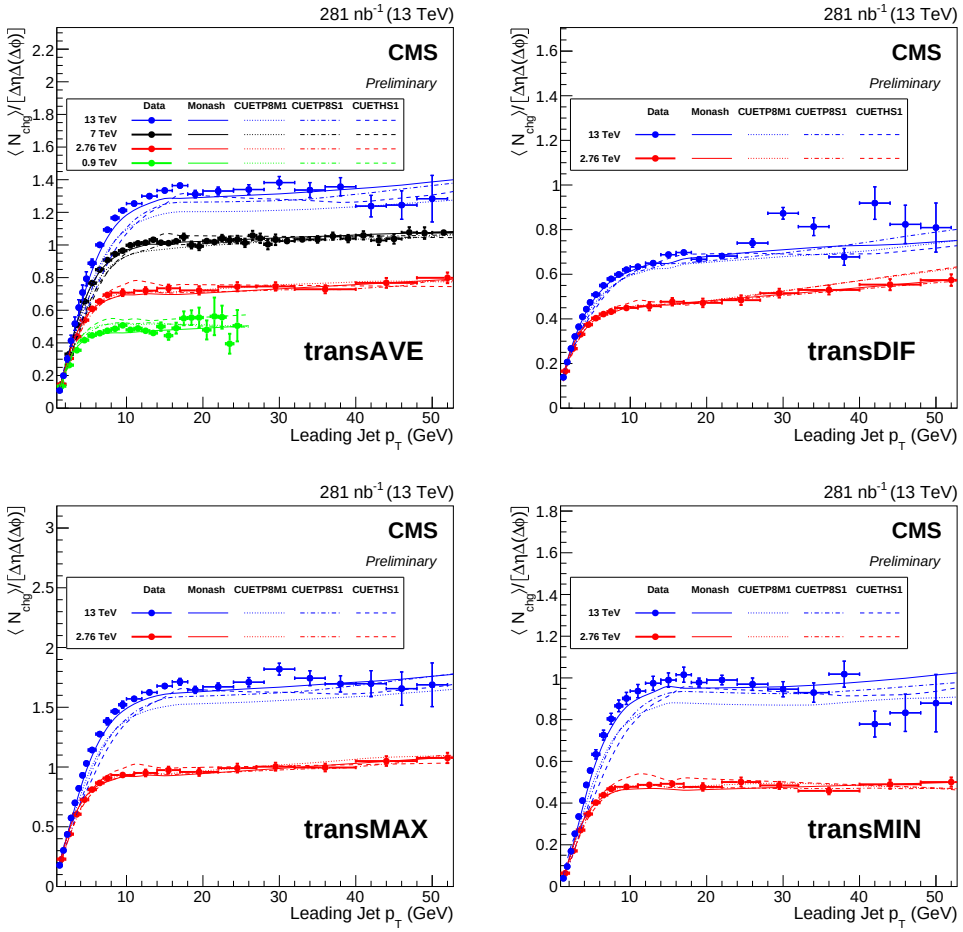

Figure 4.47 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN particle densities with various simulations at $\sqrt{s} = 0.9$, 2.76, 7, and 13 TeVfor the top left plot, and $\sqrt{s} = 2.76$ and 13 TeVfor the rest as a function of p_T^{jet}.

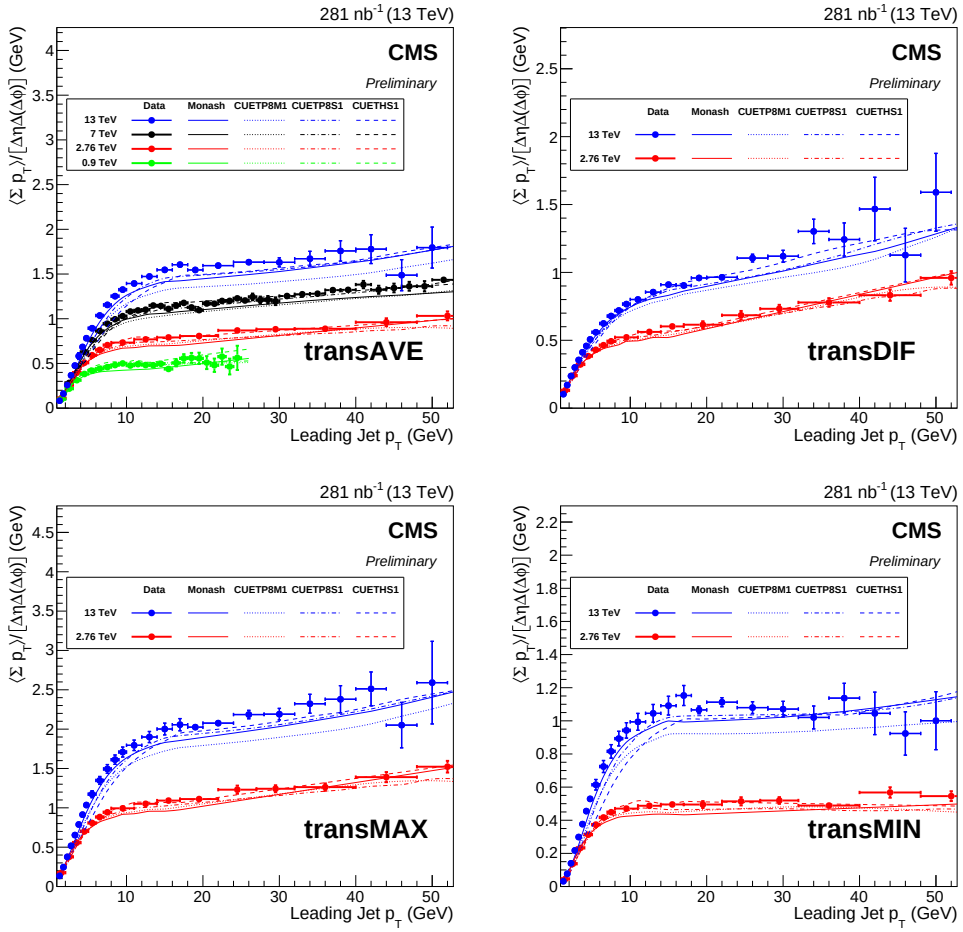

Figure 4.48 Comparisons of corrected (top left) transAVE, (top right) transDIF, (bottom left) transMAX, and (bottom right) transMIN energy densities with various simulations at $\sqrt{s} = 0.9$, 2.76, 7, and 13 TeVfor the top left plot, and $\sqrt{s} = 2.76$ and 13 TeVfor the rest as a function of p_T^{jet}.

The underlying event densities are measured as a function of p_T/p_T^{jet}, using charged-particles with $|\eta| < 2$ and $p_T > 0.5$ GeV in the region transverse to the leading charged-particle/jet. For the first time at the LHC, the transMAX, transMIN, and transDIF densities are measured, separating the various components of the UE activity. By comparing the measurements to previous results at 0.9, 2.76, and 7 TeV, and to several Monte Carlo event generator predictions, constraints are put the modelling of the underlying event. This allows the CMS Collaboration to model soft QCD better, which in turn makes measurements in searches of standard model and new physics processes more precise.

A steep rise of the UE activity in the transverse region is seen with increasing p_T/p_T^{jet}. This is followed by saturation above ~ 8 GeV, with nearly constant particle density and small Σp_T density increase. The events at the high p_T end of the distributions indicate the presence of a hard component in the transverse region. Such a distinct change in the amount of activity depending on the transverse momentum of the leading charged-particle jet is clearly seen for all the observables presented in the analyses, although for transMAX and transDIF densities, the distributions continue to show a rise with p_T/p_T^{jet} above ~ 8 GeV.

The results are compared to various tunes of PYTHIA, HERWIG++, and EPOS Monte Carlo event generators. In general, the Monte Carlo models predict the behaviour of underlying event qualitatively well, with various levels of performance between the different tunes. However it is clear that the descriptions of the data by the various predictions are much better for the transDIF densities.

By comparing data taken across $\sqrt{s} = 0.9$, 2.76, 7, and 13 TeV, a strong growth with increasing center-of-mass energy of the hadronic activity in the transverse region is also observed for the same value of the leading p_T^{jet}. A centre-of-mass energy comparison is also made for the transMAX, transMIN, and transDIF densities, showing that the transMIN densities have a stronger $\sqrt{s}$ dependence compared to the transDIF densities.

These data provide a detailed measurement of the underlying event activity in pp collisions with QCD jets, using average multiplicity and Σp_T densities. The transverse/transMAX/transMIN/transDIF densities enable more specific MC model comparisons than possible with preceding published data. These measurements are therefore expected to play a significant role in the future development and tuning of Monte Carlo event generator models of the underlying event.

Chapter 5

Conclusion

The standard model encapsulates our current knowledge of the fundamental particles that make up our universe and the electromagnetic, strong, and weak forces that governs their interaction. At the core of every element lies a nucleus composed of protons and neutrons. Since the discovery of quarks, the protons and neutrons are understood to be made up of these quarks, and the gluons that bind them together. The nuclear force that binds the protons and neutrons into the nucleus is then interpreted as some kind of residual strong force.

The physics of the residual strong force is not well understood due to the nature of strong force being unmanageably huge at large distances, resulting in confinement. For this reason, Quantum Chromodynamics, the field theoretic framework that describes the strong force works extremely well at small distance scales while becoming infinite at large distances. This directly affects the approach and limits to which we can probe the interactions between quarks and gluons. Phenomenological studies in QCD arise due to the insufficiency of perturbative methods to make definitive predictions in the very soft regime (large distance). Statistical and Monte Carlo methods thus become very important in the study of soft QCD.

The topics presented in this thesis: multiplicity distributions, angular correlations, and in particular the measurement of the underlying event, sheds light on the nature of the strong interaction in the soft regime. Statistical modelling of multiplicity distributions show that the broadening of the distribution with centre-of-mass collision energy can be better described by models which incorporate multiple initial particle emission sources. These include the Negative

Binomial Distribution, Generalised Multiplicities Distribution, and the Weighted Generalised Multiplicities Distribution. The emergence of the shoulder-like structure has however remains a challenge for these models which attempt to describe the global characteristics, including very soft particles. This is partly due to the extremely complicated and messy nature of non-abelian dynamics that give rise to confinement, and the attempt to model such processes using only a few parameters that possess a physical interpretation. On the other hand, the numerical approach yields a possible explanation of the shoulder-like structure. This is done by the introduction of Multi-Parton Interactions, implemented concurrently with perturbative Quantum Chromodynamics in Monte Carlo event generators. Such an approach describes the shoulder-like structure well, but still fails in the very high multiplicity regime.

Meanwhile the observation of the long-range near-side ridge in two-particle angular correlations in pp collisions, has resulted a controversy regarding its physical origin. Hydrodynamical models which were applied to describe the same long-range near-side ridge in heavy ion collisions, were naturally extended to describe the same phenomena in pp collisions. Other models like colour connections in the longitudinal direction, medium-jet interactions, colour glass condensates, and even Multi-Parton Interactions have also been proposed to explain the origin of the ridge.

All these experimental observations only point towards our lack of understanding of soft interactions. At the same time, these interactions have deep implications for high power, high luminosity collider physics. Not only are pp collisions dominated by the effects of soft interactions between a large number of partons with small longitudinal momentum fraction, the effect is amplified by the large number of overlapping pp collisions within a single proton bunch crossing. Thus such soft interactions have to be well understood for precise measurements of other standard model and new physics processes, since they form an irreducible background to these measurements.

In order to understand the nature of soft interactions, the measurement of the underlying event becomes increasingly important for the particle physics community. The measurement of the underlying event at 2.76 and 13 TeV thus serves to quantify the final-state radiative contribution by soft interactions, and its centre-of-mass energy dependence. Results have shown that the underlying event activity rises steeply as a function of the leading charged-particle or leading

charged-particle jet transverse momentum, followed by a saturation of radiative activity at large transverse momentum scales. This corroborates the hypothesis that Multi-Parton Interactions are involved in the production of soft radiation. At the same time, the underlying event activity exhibits a strong growth with the centre-of-mass collision energy. This hints at the increased density of small longitudinal momentum fraction partons within the colliding protons, as measured in deep inelastic scattering experiments.

The underlying event activity is qualitatively well described by various competing Monte Carlo generators. These event generators have slight variances in their implementation of the Multi-Parton Interaction model. By comparing the underlying event activity to data and simulation, effects due to different implementation can be examined, and the parameters can be optimised for describing soft radiation better.

The measurement of the underlying event activity thus improves our physical understanding of strong interactions in the very soft regime, and optimises the performance of the Monte Carlo event generator models which summarise our knowledge of fundamental particle interactions. As such, the measurement is expected to play a significant role for the CMS Collaboration and the larger particle physics community.

References

[1] W. K. H. Panofsky, "The Evolution of Particle Accelerators & Colliders," *Beam Line*, vol. 27, p. 36, 1997.

[2] C. B. Parker, *McGraw Hill Encyclopaedia of Physics (2nd ed.)*. Mc Graw Hill, 1994.

[3] M. Mansfield and C. O'Sullivan, *Understanding Physics (4th ed.)*. John Wiley & Sons, 2011.

[4] Particle Data Group, "Review of Particle Physics," *J. Phys. G*, vol. 37, p. 075021, 2010.

[5] Ali, A. and Kramer, G., "JETS and QCD: a historical review of the discovery of the quark and gluon jets and its impact on QCD," *Eur. Phys. J. H*, vol. 36, p. 245, 2011.

[6] S. Bethke, "Experimental tests of asymptotic freedom," *Nucl. Phys. B*, vol. 58, p. 351, 2007.

[7] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, "Parton distributions for the LHC," *Eur. Phys. J. C*, vol. 63, p. 189, 2009.

[8] N. Cartiglia, "Measurement of the proton-proton total, elastic, inelastic and diffractive cross sections at 2, 7, 8 and 57 TeV," *[arXiv:1305.6131]*, 2013.

[9] CMS Collaboration, "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV," *JHEP*, vol. 2011, p. 1, 2011.

[10] UA5 Collaboration, "Scaling violations in multiplicity distributions at 200 and 900 GeV," *Phys. Lett. B*, vol. 167, p. 476, 1986.

[11] ALICE Collaboration, "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 0.9$ and 2.36 TeV with ALICE at LHC," *Eur. Phys. J. C*, vol. 68, p. 89, 2010.

[12] A. K. Dash and B. Mohanty, "Extrapolation of multiplicity distribution in $p + p(\overline{p})$ collisions to LHC energies," *J. Phys. G*, vol. 37, p. 025102, 2010.

[13] UA5 Collaboration, "Scaling violation favouring high multiplicity events at 540 GeV CMS energy," *Phys. Lett. B*, vol. 138, p. 304, 1984.

[14] CMS Collaboration, "First measurement of the underlying event activity at the LHC with $\sqrt{s} = 0.9$ TeV," *Eur. Phys. J. C*, vol. 70, p. 555, 2010.

[15] CMS Collaboration, "Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV," *JHEP*, vol. 09, p. 109, 2011.

[16] N. Cabibbo, G. Parisi, and M. Testa, "Hadron production in e^+e^- collisions," *Lett. Nuovo Cim.*, vol. 4, p. 35, 1970.

[17] S. D. Drell, D. J. Levy, and T.-M. Yan, "Theory of Deep-Inelastic Lepton-Nucleon Scattering and Lepton-Pair Annihilation Processes. III. Deep-Inelastic Electron-Positron Annihilation," *Phys. Rev. D*, vol. 1, p. 1617, 1970.

[18] S. M. Berman, J. D. Bjorken, and J. B. Kogut, "Inclusive Processes at High Transverse Momentum," *Phys. Rev. D*, vol. 4, p. 3388, 1971.

[19] G. Hanson *et al.*, "Evidence for Jet Structure in Hadron Production by e^+e^- Annihilation," *Phys. Rev. Lett.*, vol. 35, p. 1609, 1975.

[20] G. Hanson *et al.*, "Hadron production by e^+e^- annihilation at center-of-mass energies between 2.6 and 7.8 GeV. II. Jet structure and related inclusive distributions," *Phys. Rev. D*, vol. 26, p. 991, 1982.

[21] R. Brandelik *et al.*, "Evidence for planar events in e^+e^- annihilation at high energies," *Phys. Lett. B*, vol. 86, p. 243, 1979.

[22] D. P. Barber *et al.*, "Discovery of Three-Jet Events and a Test of Quantum Chromodynamics at PETRA," *Phys. Rev. Lett.*, vol. 43, p. 830, 1979.

[23] C. Berger *et al.*, "Evidence for gluon bremsstrahlung in e^+e^- annihilations at high energies," *Phys. Lett. B*, vol. 86, p. 418, 1979.

[24] W. Bartel *et al.*, "Observation of planar three-jet events in e^+e^- annihilation and evidence for gluon bremsstrahlung," *Phys. Lett. B*, vol. 91, p. 142, 1980.

[25] J. Ellis, M. K. Gaillard, and G. G. Ross, "Search for gluons in e^+e^- annihilation," *Nucl. Phys. B*, vol. 111, p. 253, 1976.

[26] A. Giovannini and L. Van Hove, "Negative binomial multiplicity distributions in high energy hadron collisions," *Z. Phys. C*, vol. 30, p. 391, 1986.

[27] A. Giovannini and L. Van Hove, "Negative binomial properties and clan structure in multiplicity distributions," *A. Phys. Pol. B*, vol. 19, p. 495, 1988.

[28] A. Giovannini and L. Van Hove, "Clan Structure Analysis and QCD Parton Showers in Multiparticle Dynamics: An Intriguing Dialog Between Theory and Experiment," *Int. J. Mod. Phys. A*, vol. 20, p. 3897, 2005.

[29] ALICE Collaboration, "Charged-particle multiplicity measurement in proton-proton collisions at $\sqrt{s} = 7$ TeV with ALICE at LHC," *Eur. Phys. J. C*, vol. 68, p. 345, 2010.

[30] P. Ghosh, "Negative binomial multiplicity distribution in proton-proton collisions in limited pseudorapidity intervals at LHC up to $\sqrt{s} = 7$ TeV and the clan model," *Phys. Rev. D*, vol. 85, p. 054017, 2012.

[31] A. Giovannini, S. Lupia, and R. Ugoccioni, "Common origin of the shoulder in multiplicity distributions and of oscillations in the factorial cumulants to factorial moments ratio," *Phys. Lett. B*, vol. 374, p. 231, 1996.

[32] I. Zborovský, "A three-component description of multiplicity distributions in pp collisions at the LHC," *J. Phys. G*, vol. 40, p. 055005, 2013.

[33] J. Peřina, "Superposition of coherent and incoherent fields," *Phys. Lett. A*, vol. 24, p. 333, 1967.

[34] W. Y. Wang, Q. Leong, W. K. Ng, A. Dewanto, A. H. Chan, and C. H. Oh, "Deformed Coherent State for Multiparticle Production Mechanism," *Proceedings of the Conference in Honour of the 90th Birthday of Freeman Dyson*, p. 400, 2014.

[35] C. Chew, D. Kiang, and H. Zhou, "A generalized non-scaling multiplicity distribution ," *Phys. Lett. B*, vol. 186, p. 411, 1987.

[36] A. H. Chan and C. K. Chew, "Parton branching model for $p\overline{p}$ collisions," *Phys. Rev. D*, vol. 41, p. 851, 1990.

[37] W. Y. Wang, S. Seah, J. Setianegara, A. H. Chan, and C. H. Oh, "The weighted GMD for multiplicity distributions at LHC energies," *Proceedings of the XLV International Symposium on Multiparticle Dynamics (accepted)*, 2016.

[38] T. Sjöstrand and P. Skands, "Multiple Interactions and the Structure of Beam Remnants," *JHEP*, vol. 03, p. 053, 2004.

[39] B. Alver *et al.*, "System size dependence of cluster properties from two-particle angular correlations in $Cu+Cu$ and $Au+Au$ collisions at $\sqrt{s_{NN}} =$ 200 GeV," *Phys. Rev. C*, vol. 81, p. 024904, 2010.

[40] CMS Collaboration, "Multiplicity and transverse momentum dependence of two- and four-particle correlations in pPb and PbPb collisions," *Phys. Lett. B*, vol. 724, p. 213, 2013.

[41] B. Alver and G. Roland, "Collision-geometry fluctuations and triangular flow in heavy-ion collisions," *Phys. Rev. C*, vol. 81, p. 054905, 2010.

[42] B. Alver *et al.*, "Triangular flow in hydrodynamics and transport theory," *Phys. Rev. C*, vol. 82, p. 034913, 2010.

[43] B. Schenke, S. Jeon, and C. Gale, "Elliptic and Triangular Flow in Event-by-Event $D = 3 + 1$ Viscous Hydrodynamics," *Phys. Rev. Lett.*, vol. 106, p. 042301, 2011.

[44] H. Petersen *et al.*, "Triangular flow in event-by-event ideal hydrodynamics in $Au + Au$ collisions at $\sqrt{s_{NN}} = 200$ GeV," *Phys. Rev. C*, vol. 82, p. 041901, 2010.

[45] J. Xu and C. M. Ko, "Effects of triangular flow on di-hadron azimuthal correlations in relativistic heavy ion collisions," *Phys. Rev. C*, vol. 83, p. 021903, 2011.

[46] D. Teaney and L. Yan, "Triangularity and dipole asymmetry in relativistic heavy ion collisions," *Phys. Rev. C*, vol. 83, p. 064904, 2011.

[47] CMS Collaboration, "Observation of long-range, near-side angular correlations in proton-proton collisions at the LHC," *JHEP*, vol. 09, 2010.

[48] CMS Collaboration, "Observation of long-range, near-side angular correlations in *pPb* collisions at the LHC," *Phys. Lett. B*, vol. 718, p. 795, 2013.

[49] CMS Collaboration, "Measurement of Long-Range Near-Side Two-Particle Angular Correlations in pp Collisions at $\sqrt{s} = 13$ TeV," *Phys. Rev. Lett.*, vol. 116, p. 172302, 2016.

[50] E. Avsar *et al.*, "Eccentricity and elliptic flow in proton–proton collisions from parton evolution," *Phys. Lett. B*, vol. 702, p. 394, 2011.

[51] K. Werner, I. Karpenko, and T. Pierog, ""Ridge" in Proton-Proton Scattering at 7 TeV," *Phys. Rev. Lett.*, vol. 106, p. 122004, 2011.

[52] E. Avsar, Y. Hatta, C. Flensburg, J.-Y. Ollitrault, and T. Ueda, "Eccentricity and elliptic flow in pp collisions at the LHC," *J. Phys. G*, vol. 38, p. 124053, 2011.

[53] P. Bożek, "Collective flow in *p-Pb* and *d-Pb* collisions at TeV energies," *Phys. Rev. C*, vol. 85, p. 014911, 2012.

[54] P. Bożek and W. Broniowski, "Correlations from hydrodynamic flow in *pPb* collisions," *Phys. Lett. B*, vol. 718, p. 1557, 2013.

[55] Arbuzov, B.A., Boos, E.E., and Savrin, V.I., "CMS ridge effect at LHC as a manifestation of bremsstrahlung of gluons off quarks accelerated in a strong color field," *Eur. Phys. J. C*, vol. 71, p. 1730, 2011.

[56] K. Dusling and R. Venugopalan, "Evidence for BFKL and saturation dynamics from dihadron spectra at the LHC," *Phys. Rev. D*, vol. 87, p. 051502, 2013.

[57] K. Dusling and R. Venugopalan, "Explanation of systematics of CMS $p+Pb$ high multiplicity dihadron data at $\sqrt{s_{NN}} = 5.02$ TeV," *Phys. Rev. D*, vol. 87, p. 054014, 2013.

[58] Y. V. Kovchegov and D. E. Wertepny, "Long-range rapidity correlations in heavy–light ion collisions," *Nucl. Phys. A*, vol. 906, p. 50, 2013.

[59] C.-Y. Wong, "Momentum-kick model description of the ridge in $\Delta\varphi - \Delta\eta$ correlations in pp collisions at 7 TeV," *Phys. Rev. C*, vol. 84, p. 024901, 2011.

[60] K. Dusling and R. Venugopalan, "Comparison of the color glass condensate to dihadron correlations in proton-proton and proton-nucleus collisions," *Phys. Rev. D*, vol. 87, p. 094034, 2013.

[61] A. Dumitru, K. Dusling, F. Gelis, J. Jalilian-Marian, T. Lappi, and R. Venugopalan, "The ridge in proton–proton collisions at the LHC," *Phys. Lett. B*, vol. 697, p. 21, 2011.

[62] K. Dusling and R. Venugopalan, "Azimuthal Collimation of Long Range Rapidity Correlations by Strong Color Fields in High Multiplicity Hadron-Hadron Collisions," *Phys. Rev. Lett.*, vol. 108, p. 262001, 2012.

[63] M. Strikman, "Transverse nucleon structure and multiparton interactions," *Acta. Phys. Pol.*, vol. 42, p. 2607, 2011.

[64] S. Alderweireldt and P. Van Mechelen, "Obtaining the CMS ridge effect with Multiparton interactions," *[arXiv:1203.2048]*, 2012.

[65] CMS Collaboration, "Measurement of the underlying event activity using charged-particle jets in proton-proton collisions at $\sqrt{s} = 2.76$ TeV," *JHEP*, vol. 09, p. 137, 2015.

[66] S.Bansal, A.H.Chan, X.Janssen, C.H.Oh, N.Remortel, and W.Y.Wang, "Measurement of the Underlying Event using track-jets in p-p Collisions at 2.76 TeV," *CMS Analysis Note*, vol. CMS-AN-2012/282, 2014.

[67] CMS Collaboration, "Measurement of the underlying event activity using charged particle jets in proton-proton collisions at $\sqrt{s} = 2.76$ TeV," *CMS Physics Analysis Summary*, vol. CMS-PAS-FSQ-12-025, 2014.

[68] S.Bansal, A.H.Chan, D.Ciangottini, L.Fano, R.Field, C.H.Oh, D.Rank, and W.Y.Wang, "Underlying event via leading track and track jet at 13 TeV," *CMS Analysis Note*, vol. CMS-AN-2015/184, 2015.

[69] CMS Collaboration, "Underlying event measurements with leading particles and jets in pp collisions at $\sqrt{s} = 13$ TeV," *CMS Physics Analysis Summary*, vol. CMS-PAS-FSQ-15-007, 2015.

[70] CDF Collaboration, "Charged jet evolution and the underlying event in proton-antiproton collisions at 1.8 TeV," *Phys. Rev. D*, vol. 65, p. 092002, 2002.

[71] CDF Collaboration, "Underlying event in hard interactions at the Fermilab Tevatron $\overline{p}p$ collider," *Phys. Rev. D*, vol. 70, p. 072002, 2004.

[72] CDF Collaboration, "Studying the underlying event in Drell-Yan and high transverse momentum jet production at the Tevatron," *Phys. Rev. D*, vol. 82, p. 034001, 2010.

[73] CMS Collaboration, "Study of the underlying event at forward rapidity in pp collisions at $\sqrt{s} = 0.9$,2.76, and 7 TeV," *JHEP*, vol. 04, p. 72, 2013.

[74] ATLAS Collaboration, "Measurement of underlying event characteristics using charged particles in pp collisions at $\sqrt{s} = 900$ GeV and 7 TeV with the ATLAS detector," *Phys. Rev. D*, vol. 83, p. 112001, 2011.

[75] ALICE Collaboration, "Underlying Event measurements in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV with the ALICE experiment at the LHC," *JHEP*, vol. 7, p. 116, 2012.

[76] CMS Collaboration, "Measurement of the underlying event in the Drell-Yan process in proton-proton collisions at $\sqrt{s} = 7$ TeV," *Eur. Phys. J. C*, vol. 72, p. 2080, 2012.

[77] CMS Collaboration, "Study of the underlying event, b-quark fragmentation and hadronization properties in $t\bar{t}$ events," *CMS Physics Analysis Summary*, vol. CMS-PAS-TOP-13-007, 2013.

[78] T. Sjöstrand and M. Van Zijl, "Multiple parton-parton interactions in an impact parameter picture," *Phys. Lett. B*, vol. 188, p. 149, 1987.

[79] L. Frankfurt, M. Strikman, and C. Weiss, "Transverse nucleon structure and diagnostics of hard parton-parton processes at LHC," *Phys. Rev. D*, vol. 83, p. 054012, 2011.

[80] G. Marchesini and B. R. Webber, "Associated transverse energy in hadronic jet production," *Phys. Rev. D*, vol. 38, p. 3419, 1988.

[81] J. Pumplin, "Hard underlying event correction to inclusive jet cross sections," *Phys. Rev. D*, vol. 57, p. 5787, 1998.

[82] C.Baus *et. al.*, "Study of the Underlying Event at Forward Rapidity in Proton-Proton Collisions at the LHC," *CMS Analysis Note*, vol. CMS AN-2012/064, 2012.

[83] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "Pythia 6.4 physics and manual," *JHEP*, vol. 05, p. 026, 2006.

[84] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1," *Comput. Phys. Commun.*, vol. 178, p. 852, 2008.

[85] M. Bähr, S. Gieseke, M. A. Gigg, D. Grellscheid, K. Hamilton, O. Latunde-Dada, S. Plätzer, P. Richardson, M. H. Seymour, A. Shrestnev, and B. R. Webbers, "Herwig++ physics and manual," *Eur. Phys. J. C*, vol. 58, p. 639, 2008.

[86] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," *Phys. Rept.*, vol. 97, p. 31, 1983.

[87] G. A. Schuler and T. Sjöstrand, "Hadronic diffractive cross sections and the rise of the total cross section," *Phys. Rev. D*, vol. 49, p. 2257, 1994.

[88] A. Aktas *et al.*, "Measurement and QCD analysis of the diffractive deep-inelastic scattering cross-section at HERA," *Eur. Phys. J. C*, vol. 48, p. 715.

[89] T. Sjöstrand and M. Van Zijl, "A Multiple Interaction Model for the Event Structure in Hadron Collisions," *Phys. Rev. D*, vol. 36, p. 2019, 1987.

[90] CMS Collaboration, "Measurement of energy flow at large pseudorapidities in pp collisions at $\sqrt{s} = 0.9$ and 7 TeV," *JHEP*, vol. 11, p. 148, 2011.

[91] CMS Collaboration, "Underlying Event Tunes and Double Parton Scattering," *CMS Physics Analysis Summary*, vol. CMS-PAS-GEN-14-001, 2014.

[92] R. Corke and T. Sjöstrand, "Interleaved Parton Showers and Tuning Prospects," *JHEP*, vol. 03, p. 032, 2011.

[93] M. H. Seymour and A. Siodmok, "Constraining MPI models using σ_{eff} and recent Tevatron and LHC Underlying Event data," *JHEP*, vol. 10, p. 113, 2013.

[94] R. Field, "Early LHC Underlying Event Data - Findings and Surprises." 2010.

[95] A. Buckley, H. Hoeth, H. Lacker, H. Schulz, and J. E. von Seggern, "Systematic event generator tuning for the LHC," *Eur. Phys. J.*, vol. C65, p. 331, 2010.

[96] ATLAS Collaboration, "Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC," *New J. Phys.*, vol. 13, p. 053033, 2011.

[97] ATLAS Collaboration, "Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$ and 7 TeV in a diffractive limited phase-space measured with the ATLAS detector at the LHC and new PYTHIA 6 tune," *ATLAS Conference Note*, vol. ATLAS-CONF-2010-031, 2010.

[98] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global qcd analysis," *JHEP*, vol. 07, p. 012, 2002.

[99] A. Buckley *et.al*, "Systematic event generator tuning for the LHC," *Eur. Phys. J. C*, vol. 65, p. 331, 2010.

[100] S. Agostinelli *et al.*, "Geant4—a simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, p. 250, 2003.

[101] P.Bartalini, S.Bansal, G.Cerati, Y.Chao, Y. Chang, L.Fano, R.Field, A.Lucaroni, L.Mucibello, and M.Zakaria, "The underlying event in proton-proton collisions at 7 TeV," *CMS Analysis Note*, vol. CMS AN-2010/112, 2010.

[102] CMS Collaboration, "Performance of jet reconstruction with charged tracks only," *CMS Physics Analysis Summary*, vol. CMS-PAS-JME-08-001, 2010.

[103] G. P. Salam and G. Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm," *JHEP*, vol. 05, p. 086, 2007.

[104] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_T jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[105] G. D'Agostini, "A Multidimensional unfolding method based on Bayes' theorem," *Nucl. Instrum. Meth. A*, vol. 362, p. 487, 1995.

[106] Ferenc Sikler, "Spectra of Charged Hadrons in pp collisions at $\sqrt{s} = 0.9$, 2.76 and 7 TeV identified via tracker energy loss," *CMS Analysis Note*, vol. CMS AN-10-143, 2012.

[107] P. Skands, S. Carrazza, and J. Rojo, "Tuning PYTHIA 8.1: the Monash 2013 tune," *Eur. Phys. J. C*, vol. 74, 2014.

[108] T. Pierog and K. Werner, "EPOS model and ultra high energy cosmic rays," *Proceedings of the XV International Symposium on Very High Energy Cosmic Ray Interactions (ISVHECRI 2008)*, vol. 196, p. 102, 2009.

[109] H. J. Drescher and others, "Parton-based Gribov-Regge theory," *Phys. Rept.*, vol. 350, p. 93, 2001.