Study of the Dynamics of Soft Interactions with Two-Particle Angular Correlations at ATLAS

Cristina Oropeza Barrera

School of Physics & Astronomy
College of Science & Engineering

Submitted in fulfilment of the requirements
for the Degree of Doctor of Philosophy
at the University of Glasgow

February 2012

Abstract

Measurements of inclusive two-particle angular correlations in proton-proton collisions at centre-of-mass energies $\sqrt{s} = 900$ GeV and 7 TeV are presented. The events were collected with the ATLAS detector at the LHC, using a single-arm minimum-bias trigger, during 2009 and 2010. Correlations are measured for charged particles in the kinematic range defined by a transverse momentum $p_{\mathrm{T}} > 100$ MeV and pseudorapidity $|\eta| < 2.5$. In total, integrated luminosities of 7 $\mu\mathrm{b}^{-1}$ and 190 $\mu\mathrm{b}^{-1}$ are analysed for 900 GeV and 7 TeV data, respectively. At 900 GeV only events with a charged particle multiplicity $n_{ch} \geq 2$ are analysed whereas at 7 TeV, a second phase-space region of $n_{ch} \geq 20$, with a suppressed contribution from diffractive events, is also explored.

Data are corrected using a novel approach in which the detector effects are applied repeatedly to the observable distribution and then extrapolated to a detector effect of zero.

A complex structure in pseudorapidity and azimuth is observed for the correlation function at both collision energies. Projections of the two-dimensional correlation distributions are compared to the Monte Carlo generators PYTHIA8 and HERWIG++, as well as the AMBT2B, DW and Perugia 2011 tunes of PYTHIA6. The strength of the correlations seen in the data is not reproduced by any of the models.

Declaration

I declare that the results presented here are the product of my own work, unless explicit references are given in the text to the work of others. This thesis is the result of the research I carried out in the Experimental Particle Physics group in the School of Physics & Astronomy of the University of Glasgow between October 2008 and December 2011. It has not been submitted for any other degree at the University of Glasgow or any other institution.

I performed the measurements of the depletion depth of the silicon sensors described in chapter 5 as part of my qualification task to enter the list of authors of the ATLAS collaboration.

The original idea for the probabilistic track removal correction method, described in chapter 7, came from Dr James Monk. I was involved in its development through the initial debugging stages and its subsequent application to the two-particle correlations observable. I carried out the estimation of the uncertainties on the measurement, shown in chapter 8, and the final data distributions and their interpretation in terms of an independent cluster emission model in chapter 9. The Monte Carlo generator-level distributions compared to data in section 9.2 were produced by Dr James Monk using the Rivet MC toolkit.

Cristina Oropeza Barrera

Acknowledgements

First and foremost, I would like to thank my supervisor, Dr Craig Buttar, for his invaluable and friendly support throughout my PhD. Special thanks go to Dr James Monk and Dr Emily Nurse from UCL for all their help and advice, teaching me the true spirit of collaboration. Thanks to all the members of the Glasgow PPE group, in particular to Dr Sarah Allwood-Spiers and Dr Chris Collins-Tooth for answering my random questions and for sharing with me the love for tea. To Dr Saverio D'Auria, thanks for the interesting chats while keeping me company in the ATLAS Control Room when I was baby-sitting the SCT. I am also very grateful to Dr Samir Ferrag for all the help he gave me in moments of extreme crisis. And last but not least, to Dr Arthur Moraes for the occasional chats and useful discussions, muito obrigada.

To Dr Paul Soler and Prof George Lafferty, thank you for helping me get through my Viva in one piece and for making it a very constructive experience.

Thanks to my fellow PhD students Nick, Umar, Donny, Mike, Big-Mike, Nicolás, Kenny(!), Gavin and Will for the coffee breaks, film nights and visits to the pub. A big thank you to Gen (and Aurel) for her (his) friendship and for filling my visits to CERN with tea (cheese); to Alistair, for keeping me up-to-date with the affairs of the world; and to Danilo, meu amigo fofoqueiro.

Ahora en español...

Gracias al Consejo Nacional de Ciencia y Tecnología (CONACYT) de

México por haber financiado mis estudios.

A mamá y a papá, muchas gracias por todo lo que me han dado en la vida. Ustedes me han hecho la persona que soy ahora. Gracias también a mis hermanos por su apoyo y por su cariño. Mani y Ñañita, mis sobrinos adorados, a ustedes les doy gracias por existir y por darme tanta alegría. Lo más difícil de este doctorado ha sido sin duda el no estar con ustedes y verlos crecer.

Muchas gracias a toda la banda latinoamericana en Glasgow: mexicanos, chilenos, brasileños, colombianos, argentinos y uruguayos. A pesar de la buena vibra que recibimos en estas tierras, siempre es grato sentirse de nuevo como en casa y compartir todo aquello que nos une. Gracias por las reuniones en donde deshacíamos al mundo, para después arreglarlo, y en las que sin falta cerrábamos con un buen zapateado.

Un agradecimiento muy especial para quien siempre está conmigo, día y noche, soportando viento y marea, dándome fuerzas para seguir adelante y llenándome de mucho amor. A tí, Nacho, gracias... totales!

Contents

List of Figures viii

List of Tables xviii

Preface 1

I Theoretical Background 4

1 The Physics of Soft Interactions 5

1.1 The Standard Model of Particle Physics 5

1.1.1 Fundamental Forces 5

1.1.2 Fundamental Particles 7

1.1.3 Present and Future of the Standard Model 9

1.2 Quantum Chromodynamics 10

1.2.1 QCD Lagrangian 11

1.2.2 Asymptotic Freedom and Confinement 13

1.3 Minimum Bias 15

1.4 Monte Carlo Models for Soft Interactions 17

1.4.1 The PYTHIA Approach 18

1.4.2 Soft MPI in HERWIG++ 20

1.4.3 Minimum Bias MC Tunes 22

2 Correlations in Multi-Hadron Final States **25**

2.1 Introduction and Motivation 25

2.2 Review of Experimental Results 28

2.2.1 Independent Cluster Emission 34

2.3 Observable Definition 38

II Experiment Overview **41**

3 The ATLAS Experiment at the LHC **42**

3.1 The Large Hadron Collider 42

3.1.1 CERN Accelerator Complex 42

3.1.2 Machine Design and Current Status 44

3.1.3 The LHC Experiments 46

3.2 ATLAS - A Toroidal LHC ApparatuS 48

3.2.1 Inner Detector 50

3.2.2 Calorimetry 54

3.2.3 Muon Spectrometer 56

3.2.4 Forward Detectors 57

3.2.5 Trigger and Data Acquisition 58

3.2.6 Current Status 59

4 Data Reconstruction **61**

4.1 The ATLAS Offline Software 61

4.1.1 Detector Geometry Description 61

4.1.2 Simulation Infrastructure 62

4.1.3 Reconstruction and Analysis Preparation 65

4.2 Track Reconstruction 65

4.2.1 EDM and Track Parameterisation 66

4.2.2 Inside-Out Tracking Sequence 67
4.2.3 Outside-In Tracking Sequence 69
4.2.4 Low-p_{T} Tracking Sequence 70
4.3 Vertex Reconstruction 70

5 Depletion Depth Studies for the ATLAS SCT **72**
5.1 Charge Depletion in Silicon Sensors 72
5.2 Measurements using Cosmic Data 74
5.2.1 Track Selection 75
5.2.2 Track Depth Approach 75
5.2.3 Fit to Slope Approach 83
5.2.4 Conclusion 86

III Analysis **87**

6 Event Selection & Reconstruction Efficiencies **88**
6.1 Minimum Bias in ATLAS 88
6.1.1 MBTS 89
6.1.2 BPTX 90
6.2 Event and Track Selection Requirements 91
6.3 Reconstruction Efficiencies 94
6.3.1 Tracking 94
6.3.2 Vertexing 96
6.3.3 Triggering 97

7 Correction Procedure & Closure Tests **100**
7.1 Method 100
7.1.1 Trigger and Vertex Selection 100
7.1.2 Probabilistic Track Removal 101

7.2 Monte Carlo Studies 104
7.2.1 Closure Tests 104
7.2.2 Effect of Secondary Particles 107

8 Uncertainties **113**
8.1 Extrapolation to N= −1 113
8.2 Systematics 114
8.2.1 Non-Closure of the Correction Method 114
8.2.2 Uncertainties on the Efficiencies 117
8.2.3 Additional Checks 120

9 Results **127**
9.1 Two-Dimensional Correlation Functions 127
9.2 Projections in $\Delta\eta$ and $\Delta\phi$ 130
9.3 Cluster Model Fits 141

10 Conclusions **150**

Appendices **154**

A Third-degree Polynomial Fits **154**

B Non-Closure Uncertainties for 1D Projections **158**

C Uncertainties for the 2D Distributions **162**
C.1 Non-Closure Systematics 162
C.2 Total Uncertainties 164

List of Figures

1.1 The running of $\alpha_s(Q^2)$ 14

1.2 Cross-sections for various processes in proton-proton collisions at the LHC 16

1.3 Components of the inelastic cross-section 17

2.1 Dependence of $C_n(\eta_1, \eta_2)$ on $1/(n\text{-}1)$ and $1/2(\eta_1 + \eta_2)$ 29

2.2 Angular correlation function $C(\Delta\eta, \Delta\phi)$ for 23 and 53 GeV . . 30

2.3 Inclusive correlation function $C(\eta_1, \eta_2)$ and its "short-range" and "long-range" components 31

2.4 Two-particle correlation function in $\Delta\eta$ and $\Delta\phi$ for 7 TeV pp minimum bias and high-multiplicity events with $p_{\mathrm{T}} > 100$ MeV and $1 < p_{\mathrm{T}} < 3$ GeV 33

2.5 Integrated associated yields for the ridge region as a function of p_{T} for five multiplicity bins 34

2.6 Integrated associated yields for the ridge region as a function of multiplicity 34

2.7 $R(\Delta\eta)$ distributions at 200 and 410 GeV fitted to the cluster model 36

2.8 Effective cluster size and decay width as a function of $\sqrt{s}$ from the CMS and other experiments, as well as predictions from the D6T tune of PYTHIA6 37

2.9 $R(\Delta\eta)$ distributions at 0.9, 2.36 and 7 TeV fitted to the cluster model 37

3.1 CERN Accelerator Complex 43
3.2 Schematic diagram of an LHC dipole magnet 44
3.3 Event display of a 7 TeV collision in the ATLAS detector . . . 46
3.4 Cut-away schematic view of the ATLAS detector 49
3.5 Cut-away schematic view of the ATLAS Inner Detector 51
3.6 Layout of a quarter-section of the Inner Detector and its components 51
3.7 Cut-away schematic view of the ATLAS Calorimeters 55
3.8 Cut-away schematic view of the ATLAS Muon Spectrometer . 57
3.9 Operational status of the ATLAS detector 60
3.10 Total integrated luminosity delivered by the LHC in 2011 and data-taking efficiency of the ATLAS detector 60

4.1 Diagrammatic representation of the simulation process in the ATLAS software 63
4.2 Track parameters at the perigee 68
4.3 Inside-out track reconstruction sequence 69

5.1 Schematic representation of a p-n junction 73
5.2 Schematic representation of a (p^+, n, n^+) microstrip sensor . . 74
5.3 Schematic view of a particle crossing an SCT sensor from the plane transverse to the direction of the strips 76
5.4 Track depth distribution for cosmic-ray data with 150 V bias voltage and with the solenoid magnetic field off 78

5.5 Track depth distribution in different incidence angle slices for cosmic-ray data at 150 V bias voltage and with the solenoid magnetic field off 79
5.6 Fitted track depth distributions for cosmic-ray data with 150 and 50 V bias voltage and the solenoid magnetic field off . . . 80
5.7 Fitted track depth distributions for cosmic-ray data with the solenoid magnetic field on 81
5.8 Comparison between the measured depletion depth and the expected values for samples without the solenoid magnetic field . 82
5.9 Comparison between the measured depletion depth and the expected values for samples with the solenoid magnetic field . . 83
5.10 Average cluster size as a function of the local incidence angle φ for samples with the solenoid magnetic field on 84
5.11 Examples of linear fits to average cluster size vs incidence angle distributions for cosmic-ray data with the solenoid on 85
5.12 Fitted value of the slope as a function of the square root of the bias voltage 86

6.1 Display of a 7 TeV collision event in the ATLAS detector, which triggered the MBTS counters 89
6.2 BPTX signal produced by a proton bunch passing through ATLAS eight times 91
6.3 Track reconstruction efficiency at 900 GeV as a function of p_{T} and η for tracks with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ 96
6.4 Track reconstruction efficiency at 7 TeV as a function of p_{T} and η for tracks with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$ 96
6.5 Vertex reconstruction efficiency, as a function of n_{sel}^{BS}, at 900 GeV and 7 TeV 97

6.6 Δz_0^{BS} dependence of ϵ_{vtx} at 900 GeV 98
6.7 Δz_0^{BS} dependence of ϵ_{vtx} at 7 TeV 98
6.8 Trigger reconstruction efficiency, as a function of n_{sel}^{BS}, at 900 GeV and 7 TeV 99

7.1 Probabilistic track removal method to correct for track reconstruction inefficiencies 101
7.2 $R(\Delta\eta)$ after each iteration of the probabilistic track removal correction method 103
7.3 Example of third-degree polynomial fits to the values of $R(\Delta\eta)$ in a central and an outer bin 104
7.4 Distribution of the position of the primary vertex in the z-direction for data and the reconstructed PYTHIA6 MC09 Monte Carlo tune, before and after re-weighting 105
7.5 Comparison between the generated and corrected Monte Carlo $R(\Delta\eta)$ distributions at both energies 106
7.6 Comparison between the $R(\Delta\eta)$ obtained from full simulation and reduced truth at 7 TeV 108
7.7 Estimate of the effect of physics secondaries on $R(\Delta\eta)$ 110
7.8 Fraction of secondaries as a function of p_{T} and η 112

8.1 Difference between the truth and corrected pseudorapidity distributions as a function of $R(\Delta\eta)_{corr}$, for the PYTHIA6 MC09 tune at both energies 116
8.2 Difference between the truth and corrected pseudorapidity distributions as a function of $R(\Delta\eta)_{corr}$ for different Monte Carlo tunes at 7 TeV for $n_{ch} \geq 2$ 116

8.3 Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the tracking efficiency by its systematic uncertainty 118
8.4 Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the tracking efficiency by its systematic uncertainty in opposite directions for the barrel and end-cap regions of the detector 119
8.5 Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the vertex reconstruction efficiency by its systematic uncertainty 119
8.6 Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the trigger reconstruction efficiency by its systematic uncertainty 120
8.7 Closure tests for the $R(\Delta\eta)$ distribution when using five and four track-removal iterations 121
8.8 7 TeV data $R(\Delta\eta)$ distribution corrected using four, five and six iterations of the track removal procedure 122
8.9 Effect of using fourth degree polynomial fits in the extrapolation to the corrected value of $R(\Delta\eta)$ at 7 TeV in MC09 and data 123
8.10 Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the random number generator used in the probabilistic track removal iterations in data and the Monte Carlo tune MC09 124
8.11 Generated pseudorapidity correlation function compared to the ones corrected using the TRandom2 and TRandom3 random number generators to throw away tracks 124

8.12 Fraction of nearest neighbour tracks with a hit on the Pixel b-layer as a function of the $\Delta\mathcal{R}$ separation for 7 TeV data and non-diffractive Monte Carlo MC09 in different p_{T} bins 125

9.1 Corrected $R(\Delta\eta, \Delta\phi)$ distributions for 900 GeV and 7 TeV data and the MC09 tune 129

9.2 Regions used for the one-dimensional projections of the correlation function 130

9.3 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV 132

9.4 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$ 132

9.5 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$ 133

9.6 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV 134

9.7 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$ 134

9.8 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$ 135

9.9 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV 136

9.10 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$ 136

9.11 Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$ 137

9.12 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV. 138

9.13 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$ 138

9.14 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$ 139

9.15 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV. 140

9.16 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$ 140

9.17 Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$ 141

9.18 Fitted corrected background distribution in $\Delta\eta$, obtained by integrating $\Delta\phi$ between 0 and π, for events with $n_{ch} \geq 2$ at 7 TeV 143

9.19 Cluster fits to the corrected data pseudorapidity correlation functions at 900 GeV and 7 TeV for events with $n_{ch} \geq 2$ and $n_{ch} \geq 20$ 144

9.20 Cluster fits to the pseudorapidity correlation functions of tune 4C of PYTHIA8 at 900 GeV and 7 TeV for events with $n_{ch} \geq 2$ and $n_{ch} \geq 20$ 145

9.21 Effective cluster size for data and the different MC tunes as a function of the centre-of-mass energy 148

9.22 Cluster decay width for data and the different MC tunes as a function of the centre-of-mass energy 149

A.1 Third-degree polynomial fits to all bins of the $R(\Delta\eta)$ distribution for data at 7 TeV ($n_{ch} \geq 2$) 156
A.2 Example of third-degree polynomial fits done to 10 bins of the $R(\Delta\eta, \Delta\phi)$ distribution for data at 7 TeV ($n_{ch} \geq 2$) 157

B.1 Non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between 0 and π for 7 TeV with $n_{ch} \geq 2$ 159
B.2 Non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between 0 and $\frac{\pi}{2}$ for 7 TeV with $n_{ch} \geq 2$ 159
B.3 Non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between $\frac{\pi}{2}$ and π for 7 TeV with $n_{ch} \geq 2$ 160
B.4 Non-closure systematic uncertainty for $R(\Delta\phi)$ where $\Delta\eta$ is integrated between 0 and 2 for 7 TeV with $n_{ch} \geq 2$ 160
B.5 Non-closure systematic uncertainty for $R(\Delta\phi)$ where $\Delta\eta$ is integrated between 2 and 5 for 7 TeV with $n_{ch} \geq 2$ 161

C.1 Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 900 GeV Monte Carlo (MC09) 163
C.2 Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 7 TeV Monte Carlo (MC09) for events with $n_{ch} \geq 2$ 163
C.3 Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 7 TeV Monte Carlo (MC09) for events with $n_{ch} \geq 20$ 164
C.4 Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 900 GeV . . . 165

C.5 Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 7 TeV for events with $n_{ch} \geq 2$ 166
C.6 Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 7 TeV for events with $n_{ch} \geq 20$ 167

List of Tables

1.1 The fundamental forces and their properties 6

1.2 Properties of the gauge bosons 6

1.3 Classification and properties of the quarks 8

1.4 Classification and properties of the leptons 9

1.5 Main characteristics of the different MC tunes used in the final comparisons with data 24

5.1 Cosmic-ray data samples used in the depletion depth measurements 75

5.2 Calculated values for the depletion depth for the different bias-voltage cosmic-ray data samples 82

5.3 Fitted slope values for the different bias-voltage cosmic-ray data samples 85

6.1 MBTS trigger menus at Level 1 90

6.2 Total number of selected events and tracks in 900 GeV and 7 TeV data 94

7.1 Total number of selected events and tracks in 900 GeV and 7 TeV Monte Carlo MC09 PYTHIA6 tune 106

8.1 Uncertainties on the track reconstruction efficiency 117

9.1 Cluster properties for data in all phase-space regions 146
9.2 Cluster properties for the different MC tunes in all phase-space regions 147

B.1 Non-closure systematic uncertainties for the different 1D projections at 7 TeV ($n_{ch} \geq 2$) 158

Preface

The ATLAS experiment at the Large Hadron Collider has been recording collisions since November 2009 and thus far has collected around 5.3 fb^{-1} of proton-proton data in two years of operation. The majority of the interactions are of a soft nature, producing a large number of low-momentum particles. Unfortunately, soft interactions are one of the least understood processes in Quantum Chromodynamics. The lack of applicability of perturbative calculation methods means that these processes need to be approached via phenomenological models that must be tuned to experimental data. However, these models do not have any predictive power of how the dynamics of soft interactions change with collision energy. Measurements of the properties of these interactions at the LHC contribute greatly to the development and understanding of the underlying soft processes. A powerful observable to discriminate between models and reveal information about the mechanisms of particle production is the two-particle angular correlation function, which is the subject of this dissertation.

This thesis is divided in three parts. In part I, the theory relevant to the correlation measurement is presented. Chapter 1 describes the physics behind soft processes. It starts by introducing the Standard Model of particle physics and its gauge theories. It then focuses on Quantum Chromodynamics and the properties which give rise to non-perturbative approaches. After defining the term minimum bias, the chapter finishes with an overview of two of the main

Monte Carlo models for soft interactions and the tunes used throughout this analysis. The motivation for the study of correlations in multi-hadron final states and a review of some results from previous experiments are provided in chapter 2. In the final section of this chapter, the definition of the two-particle correlation function used in this thesis is given.

Part II includes a description of the experimental apparatus and the analysis software. After a brief introduction to the LHC and its experiments, chapter 3 gives a general overview of the ATLAS detector with particular emphasis on the Inner Detector, the component relevant to the correlations measurement. Chapter 4 begins with an introduction to the ATLAS offline software and the simulation infrastructure. The core of the chapter is a description of the track and vertex reconstruction algorithms. Work carried out between 2008 and 2009 evaluating the performance of the Semi-Conductor Tracker of the ATLAS Inner Detector, using cosmic-ray data, is presented in chapter 5. The concept of depletion in silicon sensors is first introduced. Two methods for calculating the depletion depth of the SCT sensors are then described followed by some concluding remarks.

In part III the complete correlations analysis procedure is detailed. Chapter 6 lists the event and track selection requirements. A summary of the track, vertex and trigger reconstruction efficiencies resulting from such a selection is presented. A detailed description of the correction method applied to unfold the detector effects from the true physics processes is given in chapter 7. Tests performed in Monte Carlo samples to validate the method are also included. The identified sources of uncertainty for this measurement, both statistical and systematic, are summarised in chapter 8, while in chapter 9 the final data distributions and a comparison with Monte Carlo predictions are presented. An interpretation of the correlation function in terms of an independent clus-

ter emission model is shown. Finally, chapter 10 contains the conclusions of the two-particle correlations analysis.

The work presented in chapters 7, 8 and 9 was carried out with the guidance and help of Dr James Monk and Dr Craig Buttar. In particular, the novel probabilistic track removal method presented in chapter 7 was developed in conjunction with Dr James Monk, who came up with the original idea.

Part I

Theoretical Background

Chapter 1

The Physics of Soft Interactions

1.1 The Standard Model of Particle Physics

The current understanding of the particle world is encompassed in a gauge field theory[1] called the Standard Model (SM). Completed around 1970, this theory defines the fundamental components of matter and describes their interactions [1, 2]. It has been tested over and over again and its predictions have been verified to great accuracy. As an example, the existence of the $W^{\pm}$ and Z gauge bosons (section 1.1.1) and of the top quark (section 1.1.2) was predicted by the SM years before they were experimentally observed.

1.1.1 Fundamental Forces

Every interaction in nature can be attributed to one of four fundamental forces (table 1.1): electromagnetism, gravitation, strong force and weak force. In the SM, interactions are mediated by the exchange of gauge *bosons* between the particles. A boson is a particle that has an integer spin and is described by

[1] A gauge theory is one in which the Lagrangian is invariant under local transformations.

Bose-Einstein statistics. Table 1.2 summarises the properties of the known gauge bosons.

Force	Boson	Strength (coupling)	Range [m]
Electromagnetism	photon (γ)	1/137	∞
Strong Force	gluon (g)	1	10^{-15}
Weak Force	$W^{\pm}$, Z	10^{-5}	10^{-18}
Gravitation	graviton (postulated)	10^{-38}	∞

Table 1.1: The fundamental forces and their properties [1].

Boson	Charge [e]	Spin	Mass [GeV]
γ	0	1	0
g	0	1	0
$W^{\pm}$	± 1	1	80.399 ± 0.023
Z	0	1	91.1876 ± 0.0021

Table 1.2: Properties of the gauge bosons [3].

Electromagnetism describes the interactions between electric charges and is responsible for binding electrons and nuclei into atoms, and atoms into more complex compounds. It is the result of the combination of electrostatics and magnetism put forward by Maxwell in the 1860s. Its strength is inversely proportional to the distance squared r^2 so it weakens as particles move apart from each other. The mediating gauge boson is the photon, a chargeless and massless particle of light.

Gravitation is the most familiar force in the macroscopic world. Described by Einstein's Theory of General Relativity, this force governs the interactions between masses. On large scales, it describes the motion of galaxies, stars and planets, and in a more familiar way is what pulls objects "towards the ground", with a strength proportional to $1/r^2$. In the microscopic world however, it is the weakest force and there is still no quantum field theory that successfully describes and incorporates it into the SM. A mediating spin-2

gauge boson called the graviton has been postulated but no direct experimental evidence of its existence has been found.

The **Strong Force** describes the interactions between "coloured" particles (section 1.2.1) and is responsible for binding quarks (section 1.1.2) together to form protons or neutrons, and holds nucleons together to form atomic nuclei, surpassing the electromagnetic repulsion due to same sign electric charge. Unlike gravitation or electromagnetism, the strong force has the peculiar property of increasing in strength as particles move apart from each other. The force carriers are called gluons. The gauge theory describing strong interactions will be described in more detail in section 1.2.

The **Weak Force** is associated to particle decays and radioactivity (changing a particle from one type or "flavour" to another). In the previously described interactions, the mediating bosons are massless but in the case of the weak force the three carriers, $W^{\pm}$ and Z, are relatively heavy. As a consequence, the weak force is a very short-range interaction.

1.1.2 Fundamental Particles

Twelve particles constitute the building blocks of all forms of matter in the known Universe[2]. These particles are categorised into two distinct groups: quarks and leptons. Both types of particles have half-integer spin and obey the Pauli Exclusion Principle. Described by Fermi-Dirac statistics, these particles are called *fermions*.

Quarks come in six different flavours, grouped in three families or generations (table 1.3). The first generation quarks are the lightest and the most abundant in nature since they are the basic constituents of protons and neu-

[2]Less than 5% of the energy of the Universe is associated to the known visible matter. Dark matter and dark energy are the most abundant, and yet not understood, components.

trons. The heavier quarks in the second and third generations can only be produced with high-energies such as those reached in particle accelerators or cosmic-rays. A unique property of quarks is that they have a fractional electric charge. Apart from the usual intrinsic properties of electric charge, mass and spin, quarks have *colour* which means that they can interact via the strong force, making them the only fundamental particles to experience all the fundamental interactions. This property is also responsible for the fact that quarks cluster together to form colourless *hadrons*[3]. A more detailed description of quarks and the strong interaction is given in section 1.2.

Fermion		Charge [e]	Spin	Mass [GeV]
1st Generation	up (u)	$+2/3$	$1/2$	$\sim 2.5 \times 10^{-3}$
	down (d)	$-1/3$	$1/2$	$\sim 5.0 \times 10^{-3}$
2nd Generation	charm (c)	$+2/3$	$1/2$	~ 1.29
	strange (s)	$-1/3$	$1/2$	~ 0.1
3rd Generation	top (t)	$+2/3$	$1/2$	~ 173
	bottom (b)	$-1/3$	$1/2$	~ 4.2

Table 1.3: Classification and properties of the quarks [3].

Leptons are also organised in three generations, each containing one charged and one neutral particle called *neutrino* (table 1.4). As with quarks, the first generation leptons are the most abundant and, at least in the charged lepton case, the ones with the smallest mass. Muons and taus are heavier and unstable and quickly decay into lighter particles. Charged leptons interact electromagnetically and weakly. Neutrinos have a very small mass and interact only via the short-range weak force which allows them to travel long distances without being affected. Unlike quarks, leptons do not possess colour charge and hence are observed as isolated particles.

[3]Hadrons containing three (anti)quarks are called *baryons* and if they are made of a quark-antiquark pair they are called *mesons*.

Fermion		Charge [e]	Spin	Mass [GeV]
1st Generation	electron (e)	-1	$1/2$	$\sim 0.52 \times 10^{-3}$
	electron neutrino (ν_e)	0	$1/2$	$< 2 \times 10^{-9}$
2nd Generation	muon (μ)	-1	$1/2$	~ 0.105
	muon neutrino (ν_μ)	0	$1/2$	$< 2 \times 10^{-9}$
3rd Generation	tau (τ)	-1	$1/2$	~ 1.78
	tau neutrino (ν_τ)	0	$1/2$	$< 2 \times 10^{-9}$

Table 1.4: Classification and properties of the leptons [3].

For each matter particle, there exists its corresponding *anti-particle* with the same mass but opposite electric charge and inverted quantum numbers (such as baryon number, lepton number and strangeness).

1.1.3 Present and Future of the Standard Model

The SM currently encompasses two non-Abelian[4] gauge field theories: Electroweak theory (EW) and Quantum Chromodynamics (QCD). The gauge symmetry group of local transformations is $SU(3) \times SU(2) \times U(1)$.

The EW theory was developed by Glashow [4], Salam [5] and Weinberg [6] in the 1960s and unifies the electromagnetic and weak interactions into a single symmetry group $SU(2) \times U(1)$. The concept of spontaneous symmetry breaking is introduced via the Higgs mechanism [7, 8, 9] by which the $W^{\pm}$ and Z bosons acquire mass while the photon remains massless. The Higgs boson produced by this symmetry breaking is postulated as the particle responsible for giving mass to leptons and quarks, however no value of its own mass is provided by the theory. Experimental searches for the Higgs boson in several particle accelerators have narrowed the mass region in which it could exist [10, 11, 12], but no discovery has yet been made.

QCD is the gauge theory of strong interactions. It describes how quarks

[4]Local transformations do not commute with each other.

and gluons interact with each other and how they bind into hadrons. The symmetry group is $SU(3)$ and as a result, eight gauge bosons are introduced to preserve local gauge invariance. QCD has been very successful in describing experimental data from a range of particle colliders. In the following section, a more comprehensive description of this theory is presented.

Despite its many successes, the SM is not a complete theory. In addition to the lack of a quantum theory for gravitation and to the undiscovered Higgs boson, there are a number of unanswered questions that suggest that there is new physics yet to be discovered. Just to name a few, there is the mass hierarchy of the fermions, the predominance of matter over antimatter, the mass of the neutrinos, and the fact that more than 95% of the Universe is made of dark matter and dark energy. Several theories have been put forward to try to address these issues, the most popular being Supersymmetry (SUSY). This model, which postulates that each fermion will have a boson as a superpartner and vice versa, solves the hierarchy problem encountered in the SM in which to stop the mass of the Higgs boson from becoming very large, the theory relies on fine-tuning. However, there is currently no experimental evidence to support any of the Beyond the Standard Model (BSM) theories.

1.2 Quantum Chromodynamics

By 1964 a large number of hadrons had been experimentally observed which raised the suspicion that they were not elementary particles. This led Gell-Mann [13] and Zweig [14] to formulate the quark model in which hadrons are integrated by smaller components called quarks. The quark model provided a classification scheme for hadrons based on three quark flavours (only the up, down and strange quarks were known at the time). Deep inelastic scattering (DIS) experiments at the SLAC National Accelerator Laboratory in 1968

gave solid evidence that protons were not fundamental particles. After years of experimental discoveries and extensions to the theory, the quark model became the *parton* model in which hadrons are composed of quarks and gluons, collectively known as partons.

The discussion in the following sections is based on material from [15] and [16].

1.2.1 QCD Lagrangian

The Need for Colour

The existence of spin-$\frac{3}{2}$ baryons with three same flavour quarks, like Δ^{++} (uuu) and Ω^- (sss), challenged the Pauli Exclusion Principle which states that two fermions with identical quantum numbers cannot occupy the same energy level. This dilemma was solved by introducing a new degree of freedom, colour, which can take three possible values, commonly called red, green and blue. From observations it was determined that only colour neutral states can exist in nature.

Lagrangian Density

QCD is a non-Abelian gauge theory. The group of colour transformations is $SU(3)$ which gives rise to 8 gauge bosons (gluons)[5], one for each generator $\mathbf{T}^{\mathrm{i}}$ of the group. The Lagrangian density $\mathcal{L}_{QCD}$, invariant under local transformations, is given by

$$\mathcal{L}_{QCD} = \mathcal{L}_{fermion} + \mathcal{L}_{boson} + \mathcal{L}_{gauge} + \mathcal{L}_{ghost}. \tag{1.1}$$

[5] An $SU(N)$ group has $N^2 - 1$ degrees of freedom.

Interactions between fermion fields q_a are described by the first term in equation (1.1), the Dirac Lagrangian density $\mathcal{L}_{fermion}$, with

$$\mathcal{L}_{fermion} = \sum_{flav.} \bar{q}_a (i\gamma^\mu \mathbf{D}_\mu - m)_{ab} q_b, \tag{1.2}$$

where $a, b = 1, 2, 3$ are colour indices and the summation is over all quark flavours. The covariant derivative $\mathbf{D}_\mu$ allows the quarks to interact with the gauge bosons and is of the form

$$\mathbf{D}_\mu = \partial_\mu \mathbf{I} + i g_s \mathbf{T}^{\mathrm{i}} \mathcal{A}^{\mathrm{i}}_\mu, \tag{1.3}$$

with $\mathcal{A}^{\mathrm{i}}_\mu$ the gauge boson fields ($\mathrm{i} = 1, ..., 8$). The role of g_s is discussed below.

The kinetic term for the gauge bosons $\mathcal{L}_{boson}$ is

$$\mathcal{L}_{boson} = -\frac{1}{4} F^{\mathrm{i}}_{\mu\nu} F^{\mathrm{i}\mu\nu}, \tag{1.4}$$

where $F^{\mathrm{i}}_{\mu\nu}$ is the field strength tensor derived from the gluon field $\mathcal{A}^{\mathrm{i}}_\mu$ as

$$F^{\mathrm{i}}_{\mu\nu} = \partial_\mu \mathcal{A}^{\mathrm{i}}_\nu - \partial_\nu \mathcal{A}^{\mathrm{i}}_\mu - g_s f^{\mathrm{ijk}} \mathcal{A}^{\mathrm{j}}_\mu \mathcal{A}^{\mathrm{k}}_\nu. \tag{1.5}$$

From equation (1.5) it can be seen that $F^{\mathrm{i}}_{\mu\nu}$ is not gauge invariant due to the third term $g_s f^{\mathrm{ijk}} \mathcal{A}^{\mathrm{j}}_\mu \mathcal{A}^{\mathrm{k}}_\nu$. This term introduces gauge boson self interactions which implies that, unlike photons which are chargeless, gluons are coloured and can interact with each other via the strong force.

The parameter g_s, appearing in equations (1.3) and (1.5), corresponds to the QCD gauge *coupling constant* which determines the strength of the strong interaction between coloured quanta. This coupling allows the introduction of two key properties of QCD, asymptotic freedom and confinement, both of which will be discussed in section 1.2.2.

In order to make calculations of physical quantities from the Lagrangian density, the gauge needs to be fixed so that the propagators for the massless gluon fields can be defined. The computed physical quantity turns out to be independent of the choice of gauge. Commonly used is the Feynman gauge which results in a term for the Lagrangian density given by

$$\mathcal{L}_{gauge} = -\frac{1}{2}(\partial^{\mu}\mathcal{A}^{\mathrm{i}}_{\mu}). \tag{1.6}$$

By fixing the gauge, the local invariance is lost. It can be restored by introducing complex scalar fields η^{i} obeying Fermi statistics (called Faddeev-Popov ghosts) that cancel any unphysical degrees of freedom introduced by $\mathcal{L}_{gauge}$. This extra term takes the form

$$\mathcal{L}_{ghost} = \partial_{\mu}\bar{\eta}^{\mathrm{i}}(\mathbf{D}^{\mu}_{\mathrm{ij}}\eta^{\mathrm{j}}). \tag{1.7}$$

With all these ingredients, the Feynman rules for strong interactions can be derived from the Lagrangian density in equation (1.1). A discussion of Feynman rules in QCD goes beyond the scope of this thesis, so the reader is referred to textbooks such as [15] for more detailed information.

1.2.2 Asymptotic Freedom and Confinement

The QCD coupling constant g_s, introduced in the previous section, is usually expressed in terms of the effective coupling α_s as

$$\alpha_s = \frac{g_s}{4\pi}. \tag{1.8}$$

When performing high order perturbative calculations, loop diagrams emerge that introduce ultraviolet (UV) divergences. These divergences can be removed

by applying a renormalisation at a certain scale μ. This gives rise to a correction factor to the effective coupling in terms of the squared momenta Q^2 (where $Q \gg \mu$) of all particles forming a vertex

$$\alpha_s(Q^2) = \alpha_s(\mu^2) - \alpha_s(\mu^2)^2 \beta_0 \ln\left(\frac{Q^2}{\mu^2}\right) + ... \tag{1.9}$$

The β_0 coefficient is

$$\beta_0 = \frac{11N_c - 2n_f}{12\pi} \tag{1.10}$$

where N_c corresponds to the number of colours and n_f to the number of quark flavours with masses below the momentum scale Q. Since β_0 is positive, the coefficient in front of the logarithm term in equation (1.9) is negative. As a result, the effective coupling decreases as the momentum scale increases, a phenomenon known as the *running* of the coupling constant.

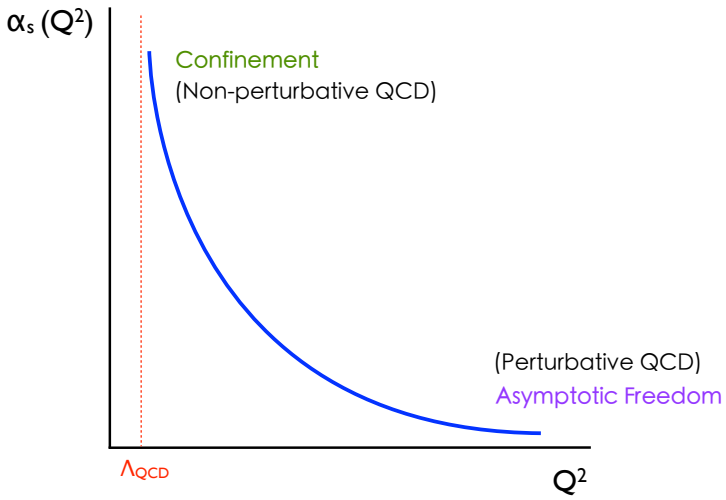


Figure 1.1: The running of $\alpha_s(Q^2)$.

Figure 1.1 illustrates the behaviour of α_s as a function of Q^2. At low momentum scales (or, equivalently, at low energies or large distances) the effective coupling is very large and so perturbative calculations cannot be performed

given that higher order terms cannot be ignored. This is the regime of *non-perturbative QCD* in which phenomenological models are needed to describe the underlying soft (low momentum transfer) physics processes. At these momentum scales, the interaction between quarks is very strong and grows even stronger when the quarks are pulled apart, until the colour string breaks forming new colourless hadrons. In summary, at macroscopic distances no isolated quarks can be observed, a property known as *confinement*. On the other hand, *asymptotic freedom* occurs at large momentum/energy scales (at short distances) when the quarks and gluons behave as quasi-free partons. In this case since α_s, the expansion parameter in the perturbative series calculations, is small the Feynman rules of *perturbative QCD* are applicable.

1.3 Minimum Bias

The total proton-proton cross-section σ_{tot} at the LHC, shown in figure 1.2, is a combination of elastic σ_{el} and inelastic σ_{inel} components [17].

In elastic events, the two protons interact from long distances and don't break apart, continuing their travel down the beam pipe. The inelastic cross-section has contributions from non-diffractive (ND) and single- (SD) and double-diffractive (DD) elements:

$$\sigma_{inel} = \sigma_{ND} + \sigma_{SD} + \sigma_{DD} \tag{1.11}$$

Diffractive events are characterised by the excitation of one (SD) or both (DD) of the protons into a high-mass colour singlet state that then decays into a shower of particles in the high rapidity direction. The non-diffractive part of the inelastic cross-section is the dominant one, and here, the two protons collide "head-on". The partons are travelling so close that they interact

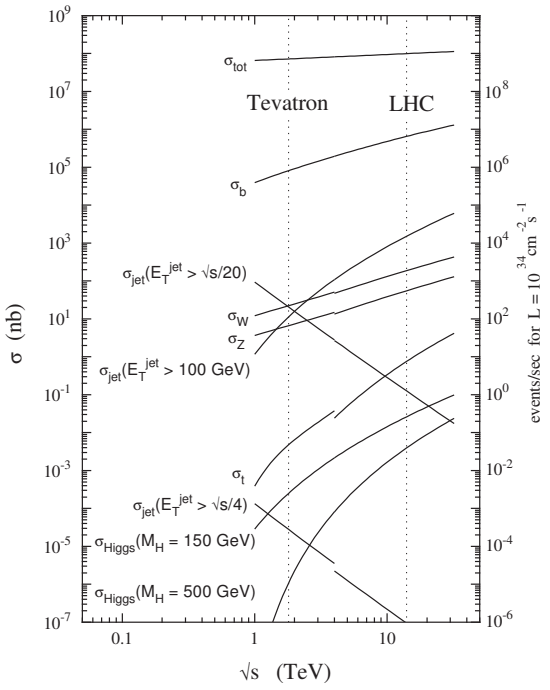


Figure 1.2: Cross-section for various processes in proton-proton collisions at the LHC. From [18].

by exchanging gluons. Most of the times, this exchange will be soft (low momentum transfer), which results in the production of low-momentum particles spread uniformly across the detector (typically soft interactions lead to low multiplicity events). On rare occasions, there could be a "hard-scatter" in which high-momentum particles and jets are produced. Figure 1.3 illustrates the different contributions to the inelastic cross-section.

The term *minimum bias* is defined experimentally and it refers to events that are selected with very loose trigger requirements, accepting a very large fraction of the inelastic cross-section. Minimum bias events are dominated by soft interactions. As such, the measurement of inclusive charged-particle distributions in minimum bias events can be used to constrain the phenomenological models of soft processes and their energy dependence. Furthermore, at high-luminosity, there will be multiple proton-proton interactions per beam crossing which means that any "interesting" hard event will most likely be

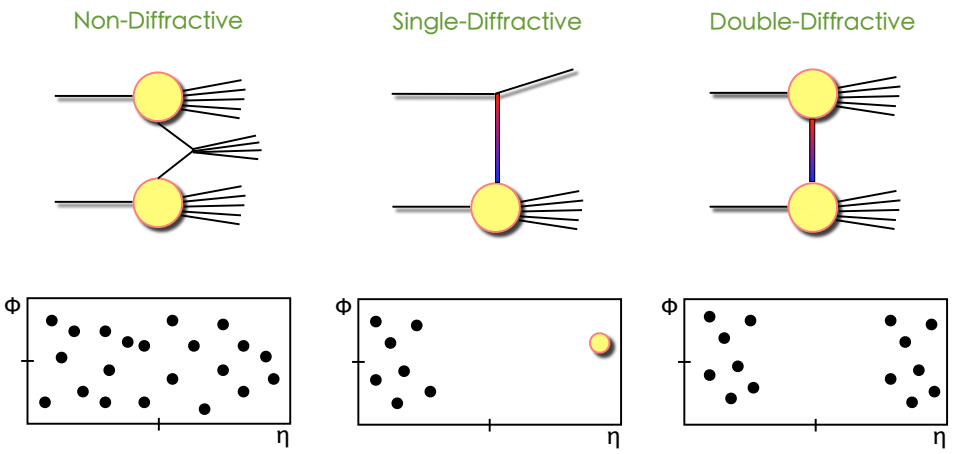


Figure 1.3: Components of the inelastic cross-section. Diagrams for each process are shown at the top and in the bottom, their corresponding signature in the $\eta - \phi$ plane of detector. For a definition of the variables η and ϕ, refer to section 2.1.

accompanied by several minimum bias events; this suggests that the characterisation and modelling of such events is key to disentangle the rare process from the very particle-dense environment.

1.4 Monte Carlo Models for Soft Interactions

As discussed in previous sections, soft parton-parton interactions are not well described by QCD because perturbative calculations are not applicable. To simulate high-energy physics events, Monte Carlo (MC) event generators implement a combination of perturbative calculations for high-p_{T}[6] processes and phenomenological models, tuned to experimental data, for low-p_{T} processes. The PYTHIA [19] and HERWIG [20] generators are used for comparisons to experimental data in the analysis presented in this thesis. A description of how these generators treat and model soft interactions within hadron collisions are

[6] p_{T} corresponds to the component of the charged-particle's momentum transverse to the direction of the incoming particles.

discussed in sections 1.4.1 and 1.4.2.

Apart from the two generators described here, there are others that implement different phenomenological models to describe soft interactions. One of them is PHOJET [21] which implements the ideas of the Dual Parton Model (DPM) to describe soft processes via Pomeron exchange [22], and perturbative QCD to generate the hard interactions.

1.4.1 The PYTHIA Approach

PYTHIA is an event generator that simulates a wide range of particle physics processes, both within and beyond the SM, with particular emphasis on those producing multi-hadron final states. PYTHIA is the most widely used generator in the field.

The event generation starts with the simulation of the "hard scatter", i.e. of the selected physics process. Once the parton scatter is sorted, PYTHIA adds in initial (ISR) and final state radiation (FSR) corrections by means of parton showers. Due to the composite nature of hadrons, the interaction between two partons leaves a remnant that is colour-connected to the hard scatter, integrating a single fragmenting system with correlated flavour, colour and kinematic properties (longitudinal and transverse momentum).

The modelling of soft interactions in PYTHIA is based on the concept of multiple parton interactions (MPI). The apparent unitarity violation introduced by the fact that the inclusive cross-section for QCD $2 \rightarrow 2$ perturbative parton scattering, σ_{int}, is greater than the total cross-section, σ_{tot}, even at a few GeV above Λ_{QCD}, is solved by allowing a hadron-hadron collision to have more than one parton interaction (sequentially ordered as a function of the transverse momentum). However, the cross-section for QCD $2 \rightarrow 2$ processes,

as a function of the p_{T}^2 scale, given by

$$\frac{d\hat{\sigma}}{dp_{\mathrm{T}}^2} \propto \frac{\alpha_s^2(p_{\mathrm{T}}^2)}{p_{\mathrm{T}}^4}, \tag{1.12}$$

diverges as $p_{\mathrm{T}} \to 0$. Inspired by perturbative QCD, PYTHIA regularises these divergences in the cross-section calculations by introducing a phenomenological correction factor in order to be able to describe to low-p_{T} parton-parton interactions. This factor can be implemented in one of two ways: a sharp cut-off or a smooth turn-off. For the former, a lower value p_{Tmin} below which no further interactions are allowed is imposed. This ensures that the cross-section vanishes completely for $p_{\mathrm{T}} < p_{\mathrm{Tmin}}$ as

$$\frac{d\hat{\sigma}}{dp_{\mathrm{T}}^2} \propto \frac{\alpha_s^2(p_{\mathrm{T}}^2)}{p_{\mathrm{T}}^4} \to \frac{\alpha_s^2(p_{\mathrm{T}}^2)}{p_{\mathrm{T}}^4}\theta(p_{\mathrm{T}} - p_{\mathrm{Tmin}}), \tag{1.13}$$

where $\theta(p_{\mathrm{T}} - p_{\mathrm{Tmin}})$ is a step function given by

$$\theta(p_{\mathrm{T}} - p_{\mathrm{Tmin}}) = \begin{cases} 0, & \text{for } p_{\mathrm{T}} < p_{\mathrm{Tmin}}, \\ 1, & \text{for } p_{\mathrm{T}} \geq p_{\mathrm{Tmin}}. \end{cases} \tag{1.14}$$

In the second approach, an energy-dependent parameter p_{T0} is introduced to smoothly regularise the divergences of the matrix elements by also modifying the dependence of the effective coupling

$$\frac{d\hat{\sigma}}{dp_{\mathrm{T}}^2} \propto \frac{\alpha_s^2(p_{\mathrm{T}}^2)}{p_{\mathrm{T}}^4} \to \frac{\alpha_s^2(p_{\mathrm{T}}^2 + p_{\mathrm{T0}}^2)}{(p_{\mathrm{T}}^2 + p_{\mathrm{T0}}^2)^2}. \tag{1.15}$$

With these changes, a continuous p_{T} spectrum is obtained. The values of p_{T0} and p_{Tmin} are two of the main free parameters of the PYTHIA model for soft processes and multiple parton interactions and need to be tuned to

experimental minimum bias and underlying event[7] data.

Given that hadrons are extended objects, PYTHIA introduces a dependence on the impact parameter (or centrality) of the interactions and on the matter distribution inside the hadrons. A small impact parameter, i.e. a large overlap of the two incoming hadrons, translates into an enhancement of the probability of having multiple interactions.

The last stage of the event generation is hadronisation in which partons are fragmented into final state hadrons. PYTHIA is based on the Lund string fragmentation model [23]. In this model, the colour energy between particles is confined in a tube-like region or *string*. As the distance between the particles increases, so does the potential energy of the string connecting them until it breaks into a new quark-antiquark pair that re-connects with the original partons to form colour-singlet systems. Further string breaks occur until only on-shell hadrons remain.

1.4.2 Soft MPI in HERWIG++

HERWIG++ is a general-purpose MC event generator of high energy collisions (lepton-lepton, lepton-hadron, hadron-hadron) that gives particular emphasis to the simulation of QCD radiation.

The event generation in HERWIG++ follows the same sequence as described in the previous section for PYTHIA. The hard process is simulated first and its energy scale and colour flow set the initial conditions for the subsequent addition of ISR and FSR. A model for multiple parton interactions allows, just as in PYTHIA, the introduction of soft interactions [24]. Including additional

[7] A complete description of a collision event requires knowledge not only of the hard-scatter process but also of contributions from multiple parton interactions, beam remnants and radiation. The underlying event is then defined as all the activity in the event but the hard process. It is an unavoidable background to most observables.

semi-hard and soft partonic scatters in HERWIG++ results in the modelling of both minimum bias and underlying event activity.

HERWIG++ uses an eikonal approximation to model semi-hard multiple interactions where $p_{\mathrm{T}} \geq p_{\mathrm{T}}^{min}$ [8]. This approach assumes that the multiple scatters are independent from each other and that the matter distribution in the hadron, as a function of the impact parameter $\mathbf{b}$, is given in terms of the inverse proton radius (one of the free parameters of the model).

In principle, an eikonal approximation works for processes in which relatively high-momentum particles scatter or when the scattering angle is small. HERWIG++ extends the concept of independent scatters into the infrared region where $0 \leq p_{\mathrm{T}} < p_{\mathrm{T}}^{min}$. The number of soft and semi-hard interactions per event is determined probabilistically. Once the evolution of the main hard scatter is completed, the remnants of the interaction are taken as the incoming hadrons for the simulation of the secondary semi-hard processes in the first place, followed by the soft interactions.

The value of p_{T}^{min}, a free parameter of the model, can be thought of as a scale at which the model makes a smooth transition from the soft non-perturbative regime to the semi-hard perturbative one.

HERWIG++ uses a cluster hadronisation model [25]. In this approach, left over gluons from the parton scatter are split into quark-antiquark pairs. Colour singlet states formed by these pairs are grouped together into clusters, which decay isotropically into the observed hadrons. Clusters are independent from each other. If the cluster is integrated by a quark-antiquark pair of flavours $(q_1, \bar{q}_2)$, a quark-antiquark (or diquark-antidiquark) pair $(q, \bar{q})$ is

[8] Here p_{T}^{min} has the same role as p_{T0} in the PYTHIA case, i.e. to stop the differential cross-section from diverging as $p_{\mathrm{T}} \to 0$. The difference is that in HERWIG++ this parameter does not have an energy-dependent evolution.

extracted from the vacuum to produce a pair of hadrons of flavours $(q_1, \bar{q})$ and $(q, \bar{q}_2)$. This is a simple and clean model for hadronisation.

1.4.3 Minimum Bias MC Tunes

Each MC generator provides a set of steerable parameters whose values can be altered to allow for different MC tunes. For comparisons with minimum bias data, three MC event generators are considered: HERWIG++ 2.5.1 (HERWIG++) [26], PYTHIA 6.4 (PYTHIA6) [19] and, its C++ updated version, PYTHIA 8.150 (PYTHIA8) [27].

The model for soft interactions used by PYTHIA6 has evolved throughout the years. Several PYTHIA6 tunes, covering the different available approaches to the description of soft processes, are used in the comparisons to the data. The original version, referred to as the *old model* [19], uses virtuality-ordered ISR and FSR showers for only the first (hardest) interaction and a non-interleaved MPI model. The DW tune [28] is used as a representative of this old model. It was tuned to underlying event and Drell-Yan data from the Tevatron and uses the CTEQ5L parton distribution functions (PDFs) [29]. The *new model* [19, 30] uses p_{T}-ordered ISR and FSR showers and a model in which MPI are interleaved with the ISR and where each parton interaction is associated with its own set of ISR and FSR. Representing this new model are the tunes: Perugia 2011 [31], MC09 [32] and AMBT2B [33]. Perugia 2011 is tuned to a wide range of Tevatron data and early LHC measurements and uses the CTEQ5L PDFs. The MC09[9] and AMBT2B tunes were produced by the ATLAS Collaboration. MC09 was tuned to minimum bias and underlying

[9]The MC09 tune is not only used for data-MC comparisons but also for systematic studies, along with the older ATLAS AMBT1 [34] and Perugia$_0$ [35] tunes of PYTHIA6, tune 1 of PYTHIA 8.130 and the PHOJET generator.

event data from the Tevatron and uses the MRSTLO* PDFs [36]. AMBT2B, on the other hand, uses the CTEQ6L1 PDFs [37] and was developed in order to describe the ATLAS minimum bias data.

PYTHIA8 adds to the new MPI model of PYTHIA6 by interleaving not only the ISR, but also the FSR. In addition, the multiple parton interactions are no longer considered independent but are allowed to interfere. An updated model for diffraction is included in PYTHIA8 that allows for harder p_{T} and particle multiplicity distributions from the single- and double-diffractive components, compared to PYTHIA6. The PYTHIA8 4C [38] tune emerges as a result of comparisons to early LHC data[10] and uses the CTEQ6L1 PDFs.

The current version of the HERWIG++ generator does not have a model for diffraction, so it has been tuned to diffraction-suppressed LHC data. Different tunes are provided for the two centre-of-mass energies, both of them using the MRST2007LO* PDFs. At 7 TeV the underlying event tune UE7-2 is used while at 900 GeV the minimum bias tune MU900-2 is implemented [39]. As discussed in the previous section, HERWIG++ has a model for hadronisation based on cluster decay, rather than the string fragmentation used by PYTHIA. It is therefore interesting to test HERWIG++ against the present measurement in order to compare the two hadronisation models and to ascertain the importance of a diffractive component in describing minimum bias observables.

The main characteristics of the chosen tunes are summarised in table 1.5. Where applicable, the MC samples used for analysis are a mix of non-diffractive and diffractive components according to the generator cross-sections.

[10] PYTHIA8 had previously been tuned to give reasonable agreement with data from the Tevatron.

Generator	Tune	Shower	PDF	Additional Info
PYTHIA 6.4	DW	virtuality-ordered	CTEQ 5L	Non-interleaved MPI model
	MC09	p_{T}-ordered	MRSTLO*	MPI interleaved with ISR
	Perugia 2011	p_{T}-ordered	CTEQ 5L	MPI interleaved with ISR
	AMBT2B	p_{T}-ordered	CTEQ6L1	MPI interleaved with ISR
PYTHIA 8.150	4C	p_{T}-ordered	CTEQ6L1	MPI interleaved with ISR and FSR and updated model for diffraction
HERWIG++ 2.5	MU900-2	angle-ordered	MRST2007LO*	MB tune used for 900 GeV only; no model for diffraction
	UE7-2	angle-ordered	MRST2007LO*	UE tune used for 7 TeV only no model for diffraction

Table 1.5: Main characteristics of the different MC tunes used in the final comparisons with data.

Chapter 2

Correlations in Multi-Hadron Final States

2.1 Introduction and Motivation

The non-perturbative nature of soft QCD processes requires the development and use of heuristic models. Monte Carlo generators implement these models and tune them to experimental data[1] (section 1.4). Single-particle distributions, such as charged-particle multiplicity distributions, are commonly used for tuning, however, these are not exhaustive descriptions of soft processes. This is the reason why the identification of new observables that are sensitive to the differences between the tunes is of particular importance. Particle correlation measurements provide useful input for discriminating between MC models and for tuning their parameters. For example, parameters controlling the modelling of multiple parton interactions, such as the minimum p_{T} or the colour flow between the hadron remnants, and the description of the

[1] Monte Carlo models were extensively tuned to data from LEP, the Large Electron-Positron collider at CERN, which operated between 1989 and 2000.

diffractive components, among others.

The study of correlations between final-state particles is also a powerful method for investigating the underlying mechanisms of particle production. Correlated activity implies that the emitted particles interacted during the course of their production.

There are a large number of ways to explore correlations in multi-hadron final state environments including, but not limited to: angular correlations, forward-backward multiplicity correlations, mass correlations, momentum correlations, Bose-Einstein correlations (between identical bosons) and charge correlations. The analysis presented in this thesis focuses on the shape and structure of angular correlations between two particles in pseudorapidity, η, and azimuthal angle, ϕ, and explores its different components using proton-proton (pp) collision data. The precise definition of the observable used in this measurement is given in section 2.3. For this analysis, ϕ is defined as the angle measured in the transverse plane around the beam axis, and η is derived from the polar angle θ (which is measured from the beam axis) as[2]

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{2.1}$$

The two-particle correlation function is a complex observable that provides information about different dynamical components of particle interactions. The existence of correlated activity in the angular phase-space can arise in many ways. For example, a fragmenting parton will radiate particles at small η values with respect to it. If the momentum of the initiating parton is large enough, when fragmenting it can produce a shower of partons (e.g.

[2]The pseudorapidity is a very useful quantity in particle physics, and is often used instead of θ due to the fact that the particle multiplicity is more or less constant as a function of η. In the massless limit, η is equal to the rapidity $y = \frac{1}{2} \ln \frac{|E+p_z|}{|E-p_z|}$.

a quark radiates a gluon which in turn liberates a $q\bar{q}$ pair, and so on) that will then hadronise into final state particles. These hadrons will be confined in a cone-shaped region or *jet* and they will be correlated in both η and ϕ. A further source of correlations is the recoil of one particle against another, i.e. travelling in opposite directions in ϕ, due to energy/momentum conservation. In the presence of back-to-back jets these correlation effects would be more pronounced. Finally, an enhancement of correlations between particles close together in η can also arise from the decay of resonances (short-lived, unstable particles) and from Bose-Einstein interference effects in which an attraction between two identical bosons (with same-sign electric charge) in the pseudorapidity space emerges.

In addition to being an interesting observable on its own, the two-particle angular correlation distribution in *pp* collisions can be used as a baseline for measurements in heavy-ion collisions. In *pp* interactions, the partons are able to escape the collision system and hadronise, whereas in nuclei collisions, the partons experience additional interactions with the hot and dense medium and the final correlation structure is modified. These measurements can then be compared to different heavy-ion scenarios such as *jet quenching*[3] and act as a key tool in characterising the dynamical evolution of the strongly interacting medium.

The study of two-particle correlations is an active field, in particular within the heavy-ion community. With the beginning of operations of the LHC, the interest on measurements in *pp* collisions has grown. In the following section, a review of some experimental correlation results from *pp* collisions is presented.

[3]Jet quenching is a phenomenon that occurs in heavy-ion collisions in which, due to the interactions of the particles with the dense medium, the measured energy of the jet is notably reduced.

This review is not, by any means, exhaustive but merely gives the flavour of the level of activity on the field and of the lessons learnt throughout the years. A compilation of correlations results is given in [40].

2.2 Review of Experimental Results

In 1975, Eggert et al. [41] published a measurement of two-particle angular correlations from *pp* collision data from the Intersecting Storage Rings (ISR) at CERN. The analysis was done at two centre-of-mass energies (23 and 53 GeV) in a sample of non-diffractive events in a narrow multiplicity range n. The pseudorapidity correlation function C_n was defined as

$$C_n(\eta_1, \eta_2) = \rho_n(\eta_1, \eta_2) - \rho_n(\eta_1)\rho_n(\eta_2), \tag{2.2}$$

where $\rho_n(\eta_1, \eta_2)$ represents the charged-pair density and $\rho_n(\eta)$ the charged-particle density. In the absence of correlation, C_n is zero everywhere by definition.

One of the key pieces of information coming out from this paper is that $(n-1)C_n$ is independent of the particle multiplicity and of $\eta_1 + \eta_2$, and depends only on the separation between the particles $\Delta\eta = \eta_1 - \eta_2$. This is illustrated in figure 2.1, where in (a) the integral of $C_n(\eta_1, \eta_2)$, over the range $|\eta_1 - \eta_2| < 0.5$, is calculated for different values of n and plotted as a function of $1/(n-1)$, while in (b) $C_n(\eta_1, \eta_2)$ is averaged over all multiplicities and plotted as a function of $1/2(\eta_1 + \eta_2)$. Extending these ideas to the azimuthal angle ϕ, a multiplicity independent correlation function in η and ϕ can then be written as

$$C(\Delta\eta, \Delta\phi) = \langle (n-1)C_n(\eta_1, \phi_1, \eta_2, \phi_2) \rangle. \tag{2.3}$$

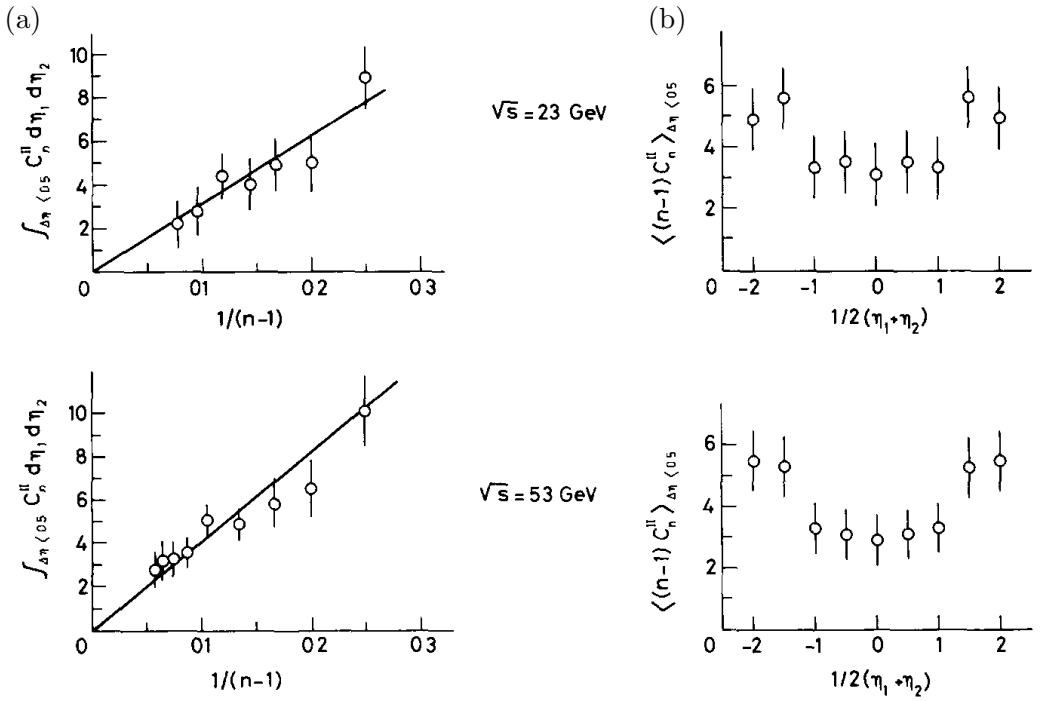


Figure 2.1: Dependence of $C_n(\eta_1, \eta_2)$ on (a) $1/(n\text{-}1)$ (b) $1/2(\eta_1 + \eta_2)$ for 23 GeV and 53 GeV data. From [41].

The relation between pseudorapidity and azimuthal correlations at these ISR energies is shown in figure 2.2. From this complex structure, three features stand out: first, a short-range (small $\Delta\eta$) peak that extends over the full $\Delta\phi$ range, second, a long-range pseudorapidity correlation around $\Delta\phi = \pi$, and third, a suppression of the correlation for large $\Delta\eta$ and small $\Delta\phi$, which decreases with energy. This means that the correlation between particles is strongest when the pair is produced in the same or opposite directions in transverse momentum.

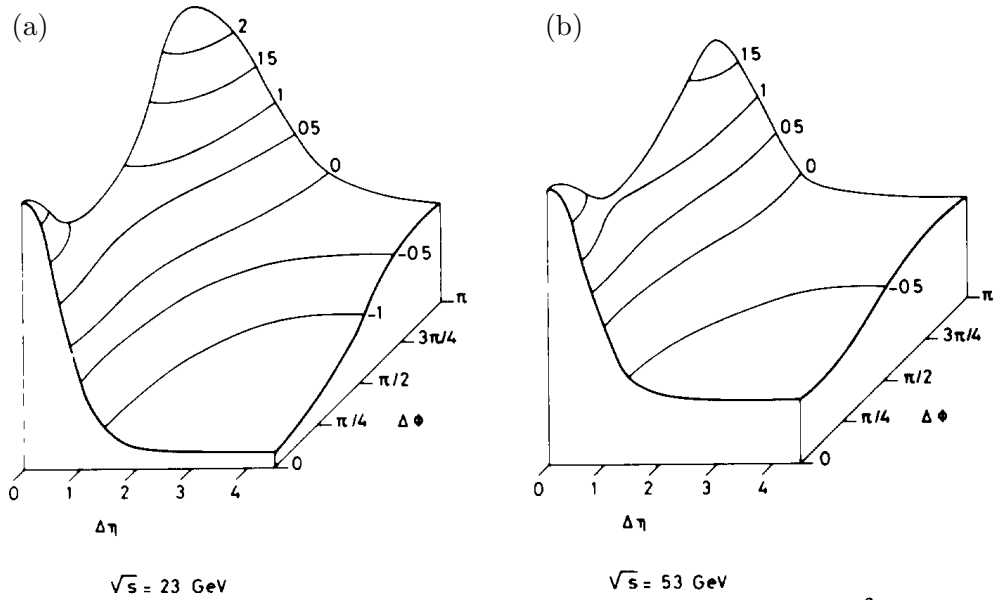


Figure 2.2: Angular correlation function $C(\Delta\eta, \Delta\phi)$ for (a) 23 GeV and (b) 53 GeV. From [41].

The UA5 Collaboration from the CERN proton-antiproton ($p\bar{p}$) collider performed an inclusive measurement (averaging over all multiplicities) of the two-particle pseudorapidity correlation function at $\sqrt{s} = 200$, 546 and 900 GeV [42]. Following the same definition for the correlation function as in equation (2.2), this collaboration found that two components contribute to the shape of the correlation function $C(\eta_1, \eta_2)$: the intrinsic correlations inside events of a given multiplicity, C_S, and the correlations arising from mixing events with different multiplicities, C_L, such that

$$C(\eta_1, \eta_2) = C_S(\eta_1, \eta_2) + C_L(\eta_1, \eta_2). \tag{2.4}$$

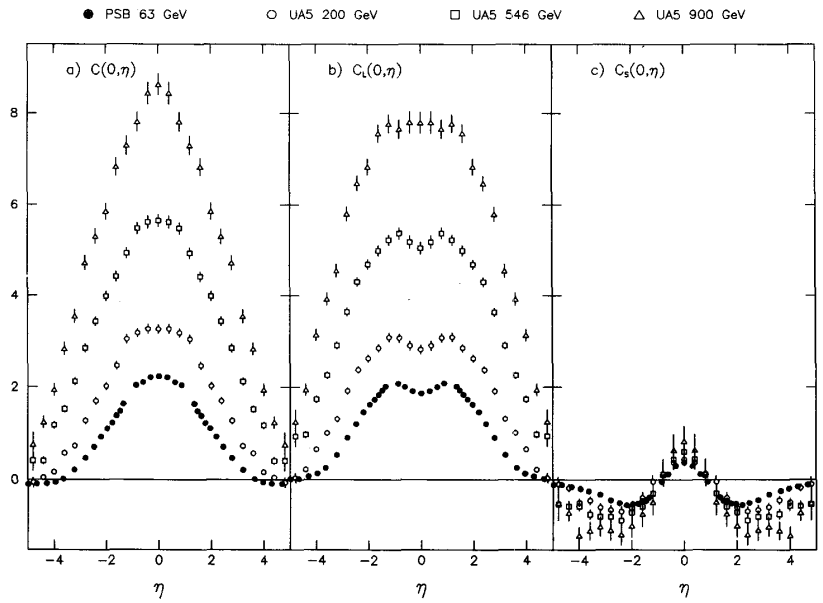


Figure 2.3: (a) Inclusive correlation function $C(\eta_1, \eta_2)$ plotted for a fixed value of $\eta_1 = 0$ versus η_2 at different c.o.m. energies and its (b) "short-range" and (c) "long-range" components. From [42].

The inclusive $C(\eta_1, \eta_2)$ is plotted in figure 2.3a for a fixed value of $\eta_1 = 0$ as a function of η_2. The first term in equation (2.4), the semi-inclusive correlation C_S (figure 2.3c), peaks at $\eta_1 = \eta_2 = 0$ and has a width of approximately 2 units in pseudorapidity, which is why it is often called "short-range" correlation. C_S does not vary much with the centre-of-mass energy. On the other hand, C_L (figure 2.3b) is present even in the absence of true dynamical correlations broadening the correlation function, so it is called "long-range" correlation. This term is responsible for the increase of the inclusive correlation with energy.

Moving forward to the present day to more up-to-date results, the CMS experiment at the LHC (section 3.1.3) has published results on two-particle correlations in pp collisions at $\sqrt{s} = 0.9$, 2.36 and 7 TeV [43]. A normalised[4]

[4]The advantage of using a normalised correlation function, compared to equation (2.2),

version of the correlation function is used, given by

$$R(\Delta\eta, \Delta\phi) = \left\langle (\langle N \rangle - 1) \left(\frac{S_N(\Delta\eta, \Delta\phi)}{B_N(\Delta\eta, \Delta\phi)} - 1 \right) \right\rangle_{bins}, \tag{2.5}$$

where the signal $S_N(\Delta\eta, \Delta\phi)$ is defined as the pair density function, the background $B_N(\Delta\eta, \Delta\phi)$ corresponds to the distribution of uncorrelated pairs, represented as a product of two single particle distributions, and $\langle N \rangle$ denotes the average particle multiplicity. The data samples are divided in bins of particle multiplicity and the correlation function is calculated for each bin and then averaged over all bins.

In addition to an inclusive analysis of the data, CMS also explored the structure of the correlation function in high-multiplicity events ($N > 110$) for low and intermediate p_{T} at 7 TeV. As shown in figure 2.4d, when the particles have an intermediate p_{T} between 1 and 3 GeV, a ridge-like structure emerges in the kinematic region $2.0 < |\Delta\eta| < 4.8$ around $\Delta\phi = 0$ in high-multiplicity events that is clearly absent from the inclusive distributions (figures 2.4a and 2.4b) and from the low-p_{T} high-multiplicity distribution (figure 2.4c). These long-range ($|\Delta\eta| < 5$) near-side ($\Delta\phi = 0$) correlations, normally observed only in heavy-ion collisions, came as a surprise and have created a lot of interest from the community as there is as yet no model that can explain them. More recently, the CMS collaboration studied the dependence of the ridge effect on the particle multiplicity and transverse momentum using their full 2010 *pp* data sample [44]. As illustrated in figure 2.5, for multiplicities above 90, the effect is zero at low-p_{T}, reaches a maximum at ∼2-3 GeV and then drops at high-p_{T}. In the p_{T} range where the effect is maximum, the ridge effect turns on around $N \sim 50$ and, with the current statistics, reaches a saturation point

is a reduced sensitivity to problems related to the acceptance of the detector.

at $N \sim 120$ as shown in figure 2.6. In the ATLAS collaboration, a search for this ridge structure in the 2010 minimum bias data was attempted, however the total number of high-multiplicity events was very low because, unlike CMS, a specialised high-multiplicity trigger was not in place when the data was collected. The high luminosity at which the LHC is currently operating makes it extremely hard for ATLAS to do a measurement of this kind. Low-luminosity runs at $\sqrt{s} = 8$ TeV are being proposed for 2012-2013 which could provide ATLAS with a window of opportunity to try to find this ridge.

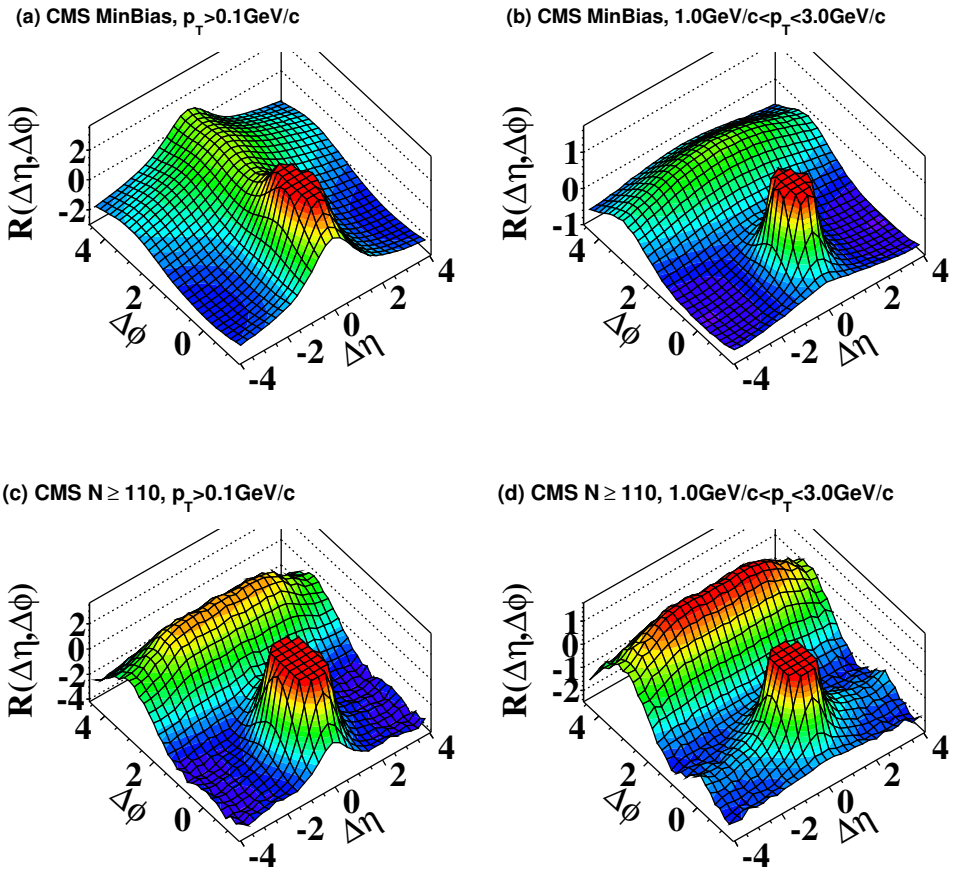


Figure 2.4: Two-particle correlation function in $\Delta\eta$ and $\Delta\phi$ for 7 TeV pp (a) minimum bias events with $p_{\mathrm{T}} > 100$ MeV, (b) minimum bias events with $1 < p_{\mathrm{T}} < 3$ GeV, (c) high-multiplicity events with $p_{\mathrm{T}} > 100$ MeV, and (d) high-multiplicity events with $1 < p_{\mathrm{T}} < 3$ GeV. The peak at (0,0) is cut off in order to reveal the underlying structure. From [43].

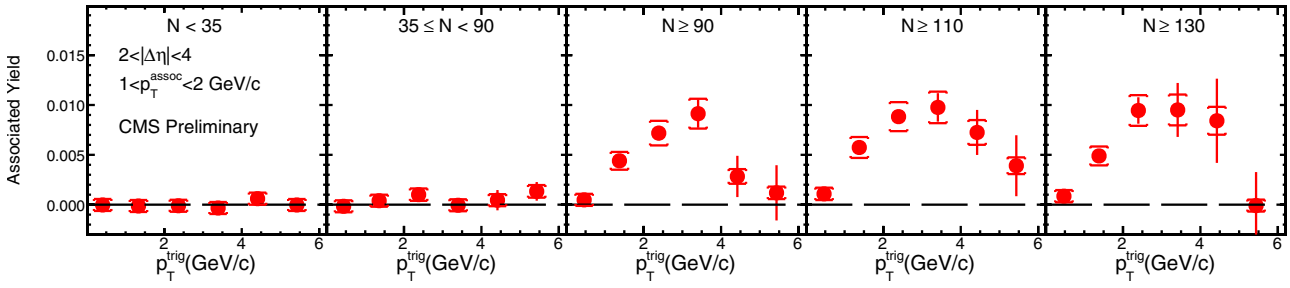


Figure 2.5: Integrated associated yields for the ridge region as a function of p_{T} for five multiplicity bins. From [44].

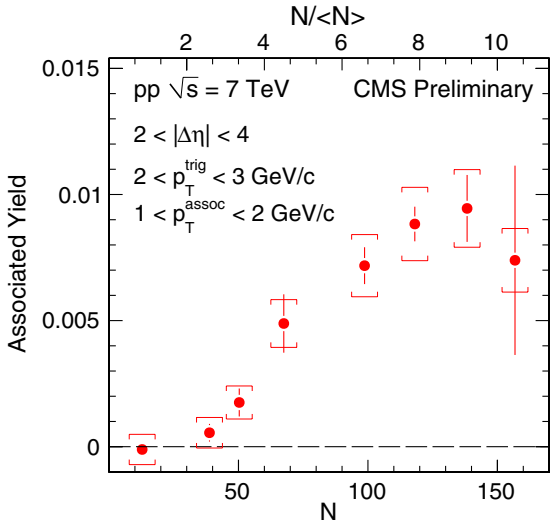


Figure 2.6: Integrated associated yields for the ridge region as a function of multiplicity for pairs where the trigger particle has a p_{T} between 2 and 3 GeV and the associated particle is in the range $1 < p_{\mathrm{T}}^{assoc} < 2$ GeV. From [44].

2.2.1 Independent Cluster Emission

The short-range nature of the pseudorapidity correlation function has often been interpreted in terms of independent cluster emission models (ICEM). The basic idea of these phenomenological models is that the observed final-state hadrons are the result of the decay of clusters. Particles originating from the same cluster will be correlated, that is, their separation in η and ϕ will be peaked around zero. On the other hand, clusters themselves are not correlated and so neither are particles from different clusters. The number

of particles may vary from cluster to cluster and the number of clusters may vary event-by-event.

A parameterisation of the correlation function $C(\Delta\eta)$ in terms of the cluster model, based on [45], can be written as

$$C(\Delta\eta) = \alpha[\Gamma(\Delta\eta) - \rho(\eta_1)\rho(\eta_2)], \tag{2.6}$$

where α denotes the strength of the correlation and represents the relative number of pairs with both particles in the same cluster, $\Gamma(\Delta\eta)$ is the pair density for particles in the same cluster and $\rho(\eta)$ corresponds to the single particle density.

The PHOBOS experiment at the Brookhaven National Laboratory's Relativistic Heavy Ion Collider investigated the cluster properties from two-particle angular correlations in pp collisions at 200 and 410 GeV [46]. Using the normalised version of equation (2.6), given by

$$R(\Delta\eta) = \alpha\left[\frac{\Gamma(\Delta\eta)}{B(\Delta\eta)} - 1\right] \tag{2.7}$$

with $B(\Delta\eta)$ the background distribution integrated over $\Delta\phi$ (equivalent to the product of the two single-particle distributions $\rho(\eta_1)\rho(\eta_2)$), two important cluster parameters can be obtained. If $\Gamma(\Delta\eta)$ is assumed to have a Gaussian shape

$$\Gamma(\Delta\eta) \propto \exp\left[-\frac{(\Delta\eta)^2}{4\delta^2}\right], \tag{2.8}$$

then the decay width of the clusters, i.e. the separation in η between the decay products of the clusters, is given by the parameter δ. The effective cluster size K_{eff} is related to α via

$$K_{\text{eff}} = \alpha + 1 = \frac{\langle K(K-1)\rangle}{\langle K\rangle} + 1. \tag{2.9}$$

The average number of particles per cluster $\langle K \rangle$ cannot be determined without any prior knowledge of the distribution of K. However, by fitting equation (2.7) to the measured pseudorapidity correlation function, K_{eff} and δ can be estimated. These fits are shown in figure 2.7.

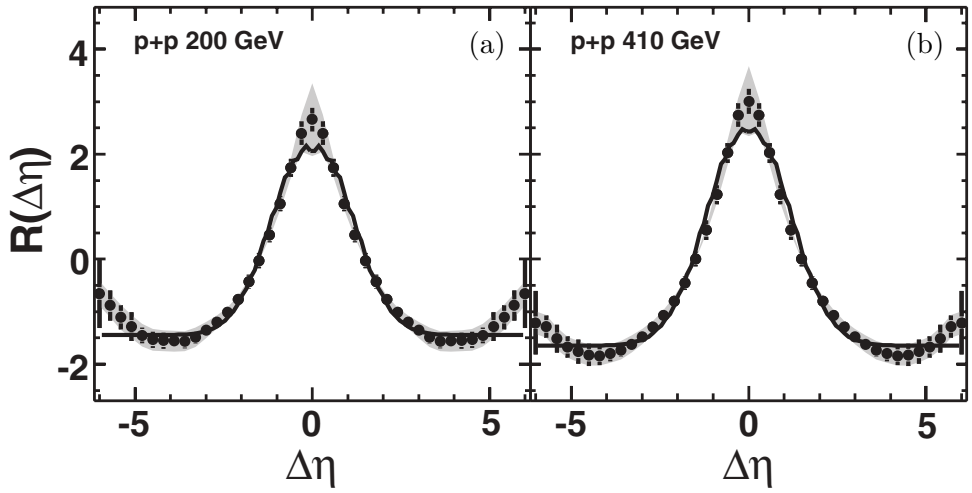


Figure 2.7: $R(\Delta\eta)$ obtained by averaging $\Delta\phi$ between 0 and π in pp collisions at (a) 200 and (b) 410 GeV. The solid curves correspond to the fits to the cluster model using equation (2.7). From [46].

The CMS collaboration published in [43] their results on cluster properties for minimum bias events as a function of the centre-of-mass energy (for particles with $p_{\text{T}} > 100$ MeV and $|\eta| < 2.4$), as well as a comparison to previous experiments, after extrapolating to their kinematic region, and to PYTHIA6 (figure 2.8). The fitted distributions are shown in figure 2.9. The overall conclusion was that the effective cluster size rises with energy, while the decay width stays approximately constant. PYTHIA6 showed the same energy dependence but underestimated the magnitude of K_{eff}.

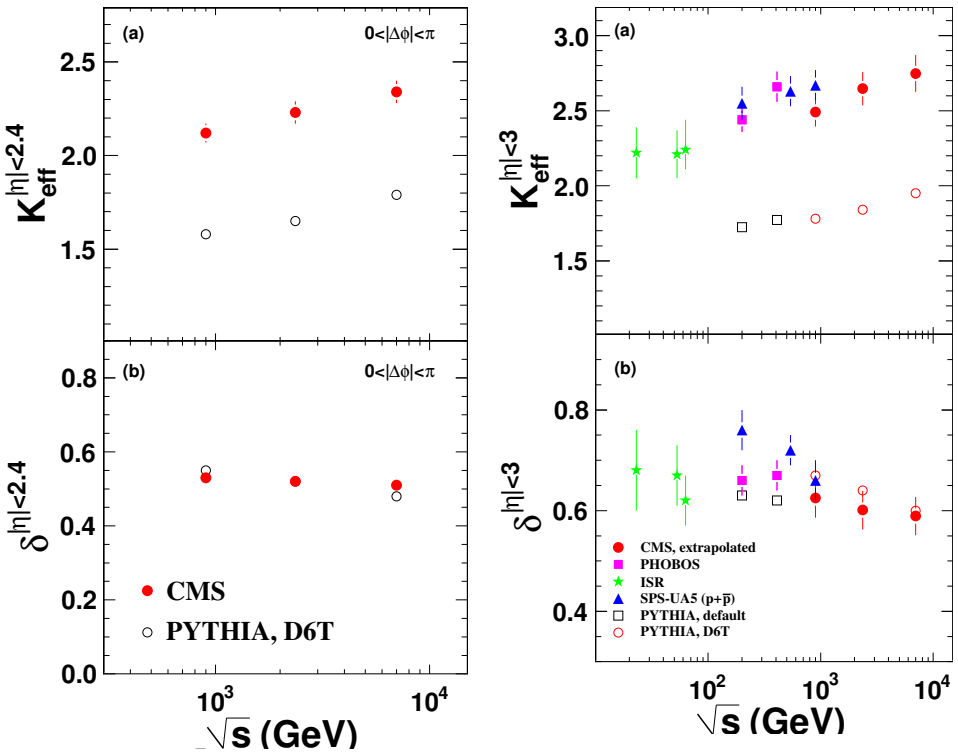


Figure 2.8: Left: (a) effective cluster size and (b) decay width as a function of $\sqrt{s}$ from the CMS experiment for particles with $p_T > 100$ MeV and $|\eta| < 2.4$. Right: comparison of the CMS (a) effective cluster size and (b) decay width results to the PHOBOS, UA5 and ISR experiments. Open circles show the predictions from the D6T tune of PYTHIA6. From [43].

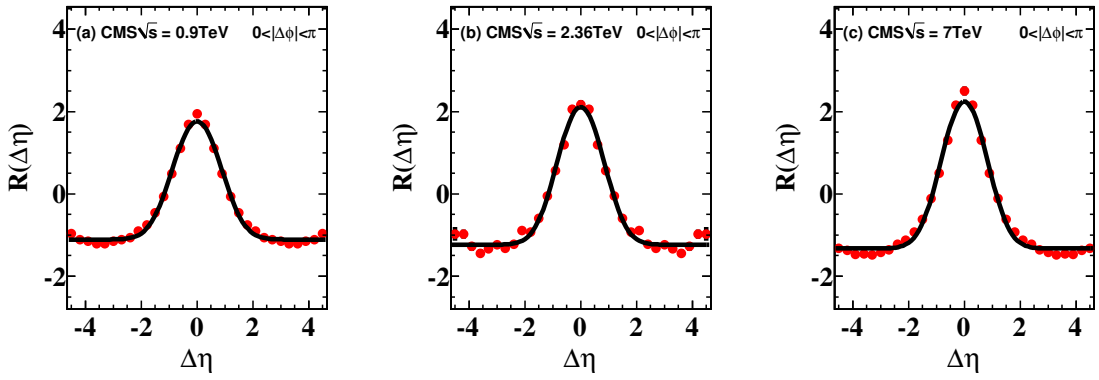


Figure 2.9: $R(\Delta\eta)$ distributions from the CMS experiment at (a) 0.9, (b) 2.36 and (c) 7 TeV. The solid curves correspond to the fits to the cluster model using equation (2.7). From [43].

2.3 Observable Definition

In order to construct the two-particle angular correlation function, two distributions in $\Delta\eta$ and $\Delta\phi$ are defined. The first distribution is the foreground $F(\Delta\eta, \Delta\phi)$, which describes the angular separation between pairs of particles emitted in the same event. It includes correlations associated with both the underlying physics process ("true correlations") and detector effects such as limited acceptance. The expression for $F(\Delta\eta, \Delta\phi)$ at a given charged-particle multiplicity n_{ch} is given by

$$F(n_{ch}, \Delta\eta, \Delta\phi) = \left\langle \frac{2}{n_{ch}(n_{ch}-1)} \sum_{i} \sum_{i \neq j} \delta_{\eta_i - \eta_j - \Delta\eta} \delta_{\phi_i - \phi_j - \Delta\phi} \right\rangle_{ch}, \qquad (2.10)$$

where the summation is over all charged particles in a single event and the average $\langle ... \rangle_{ch}$ is taken over the ensemble of events containing n_{ch} particles. The delta functions, $\delta_{\eta_i - \eta_j - \Delta\eta}$ and $\delta_{\phi_i - \phi_j - \Delta\phi}$, select particle pairs with the $\Delta\eta$ and $\Delta\phi$ separation that is appropriate for $F(\Delta\eta, \Delta\phi)$. The normalisation factor $2/n_{ch}(n_{ch}-1)$ corresponds to the inverse of the total number of charged-particle pairs in an event with n_{ch} particles. If the foreground was to be calculated using events with different multiplicities, this factor would account for the fact that a particle in a low multiplicity event would carry a lower weight than an otherwise identical particle in a higher multiplicity event because of fewer possible pair combinations.

The charged-particle density distribution, $dn_{ch}/d\eta$, is approximately flat in η. However, since the detector has a limited acceptance $\pm\eta_{max}$, phase-space alone dictates that values close to zero will be favoured for $\Delta\eta$. As such, $\Delta\eta$ is peaked at $\Delta\eta = 0$ and falls approximately linearly to a maximum possible value of $\Delta\eta = \pm 2\eta_{max}$. This motivates the second distribution used in constructing the correlation function, the background correlation B, the

$\Delta\eta$ dependence of which, for a given charged-particle multiplicity, is given by

$$B(n_{ch}, \Delta\eta) = \frac{1}{n_{ch}^2} \int_{-\eta_{max}}^{+\eta_{max}} \int_{-\eta_{max}}^{+\eta_{max}} d\eta_1 d\eta_2 \delta(\eta_1 - \eta_2 - \Delta\eta) \frac{dn_{ch}}{d\eta_1} \frac{dn_{ch}}{d\eta_2}. \quad (2.11)$$

The $\Delta\phi$ dependence of B is defined in a similar way as equation (2.11) and is flatter than the $\Delta\eta$ dependence.

With these two ingredients, the correlation function at fixed multiplicity $C(n_{ch}, \Delta\eta, \Delta\phi)$ can be constructed

$$C(n_{ch}, \Delta\eta, \Delta\phi) = F(n_{ch}, \Delta\eta, \Delta\phi) - B(n_{ch}, \Delta\eta, \Delta\phi). \quad (2.12)$$

Equation (2.12) measures whether the independent production of two particles differs from the joint production of a pair of particles with the same (η,ϕ) values. In the absence of correlation, $C(n_{ch}, \Delta\eta, \Delta\phi)$ is zero by definition.

As discussed in section 2.2, past analyses [41] have found that $(n_{ch} - 1) \times C(n_{ch}, \Delta\eta, \Delta\phi)$ is approximately independent of the particle multiplicity. Averaging $(n_{ch} - 1) \times C(n_{ch}, \Delta\eta, \Delta\phi)$ over all particle multiplicities and dividing out the background contribution gives the two-particle correlation function

$$R(\Delta\eta, \Delta\phi) = \frac{\langle (n_{ch} - 1) F(n_{ch}, \Delta\eta, \Delta\phi) \rangle}{\langle B(n_{ch}, \Delta\eta, \Delta\phi) \rangle} - \frac{\langle (n_{ch} - 1) B(n_{ch}, \Delta\eta, \Delta\phi) \rangle}{\langle B(n_{ch}, \Delta\eta, \Delta\phi) \rangle}, \quad (2.13)$$

where $\langle ... \rangle$ indicates an average over contributions from all events with all particle multiplicities. Since $dn_{ch}/d\eta$, and hence the background, does not strongly depend on the charged-particle multiplicity, the factor of B approximately cancels from the second term on the right of equation (2.13). The small multiplicity dependence of the background comes mainly from low multiplicity events where the contributions from double-diffraction tend to favour higher $\Delta\eta$ values compared to the non-diffractive component. However, low

multiplicity events are also more sensitive to track reconstruction inefficiencies because a small number of mis-reconstructed tracks has a proportionally larger impact on the event. Calculating a multiplicity independent background therefore has the advantage of diluting the effect of the experimentally more troublesome low-multiplicity events. This approach reduces the sensitivity of the observable to diffraction, which is not the motivation for this measurement.

The final expression used for the inclusive two-particle angular correlation function is given by

$$R(\Delta\eta, \Delta\phi) = \frac{\langle (n_{ch} - 1) F(n_{ch}, \Delta\eta, \Delta\phi) \rangle}{B(\Delta\eta, \Delta\phi)} - \langle n_{ch} - 1 \rangle. \tag{2.14}$$

In practice, $\langle (n_{ch} - 1) F(n_{ch}, \Delta\eta, \Delta\phi) \rangle$ is constructed by taking each pair of particles within a single event, calculating their separations in η and ϕ and filling a two-dimensional distribution with those values with a weight of $2/n_{ch}$. It is normalised by dividing each bin by the number of events entering the distribution. $B(\Delta\eta, \Delta\phi)$ is determined by taking pairs from unrelated events and, for each particle in one event, the $\Delta\eta$ and $\Delta\phi$ values with each particle in the other event are calculated and used to fill another two-dimensional distribution, which is then normalised to unit integral.

Projections of the two-dimensional correlation function along both $\Delta\eta$ and $\Delta\phi$ help reveal more details of the structure of the correlations and allow for easier comparisons with different models. These projections are calculated by first integrating separately the foreground and the background distributions over the desired ranges before taking the ratio between the two and normalising with the average particle multiplicity.

Part II

Experiment Overview

Chapter 3

The ATLAS Experiment at the LHC

3.1 The Large Hadron Collider

The European Organisation for Nuclear Research (CERN) is home to the Large Hadron Collider (LHC) [47], the biggest, most energetic particle accelerator in the world. It is installed in the existing 26.7 km tunnel of its predecessor, the Large Electron-Positron collider (LEP), which is located between 45 and 170 metres below ground, spanning across France and Switzerland. The LHC is designed to collide protons and heavy ion nuclei at very high energies and intensities in its quest to increase our understanding of nature by revealing physics processes beyond the Standard Model.

3.1.1 CERN Accelerator Complex

Before particles are injected into the LHC they have to go through a series of accelerators, each boosting the energy of the beams to increasingly higher values. Figure 3.1 shows the CERN Accelerator Complex in full detail.

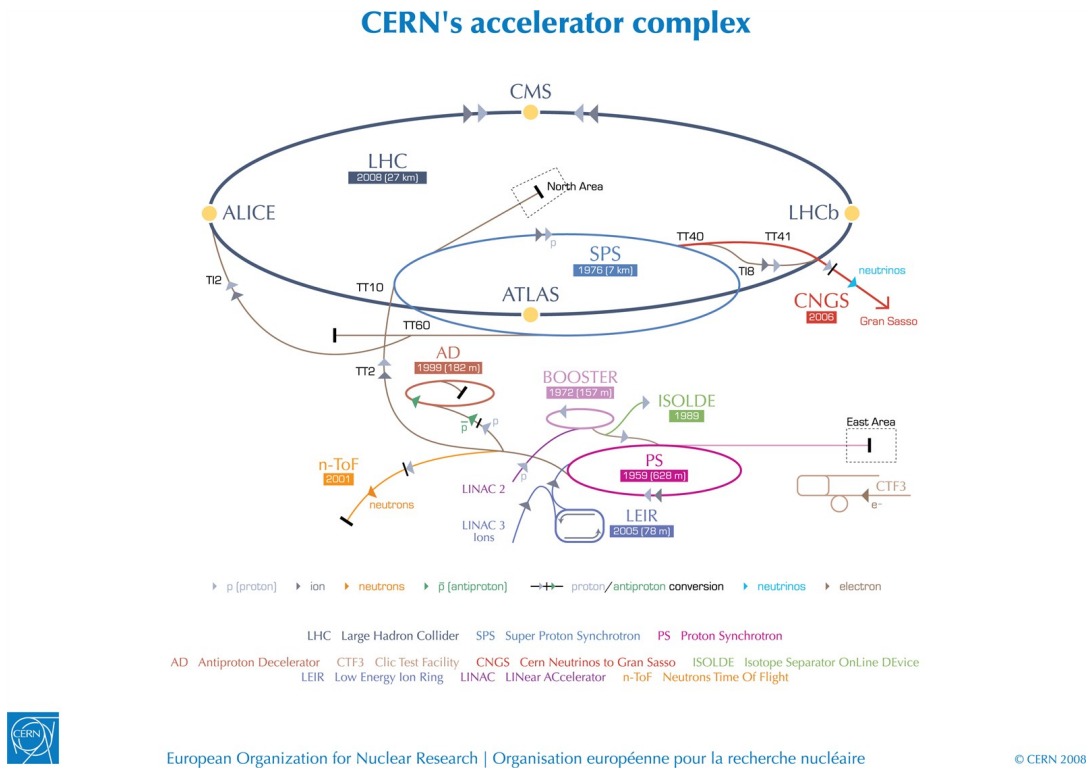


Figure 3.1: CERN Accelerator Complex. From [48].

For proton physics, 50 MeV particles (obtained by ionising hydrogen) are transferred from a Linear Accelerator LINAC2 to the Proton Synchrotron Booster where they reach an energy of 1.4 GeV. They are then injected into the Proton Synchrotron (PS), which can operate up to 25 GeV, and finally into the Super Proton Synchrotron (SPS) where they receive their final energy *kick* to 450 GeV before entering the LHC rings.

In the case of heavy-ion runs, 4.2 MeV/u lead nuclei (obtained from ionised highly-purified lead gas) move from the Linear Accelerator LINAC3 into a Low Energy Ion Ring (LEIR) where they are accelerated to 72 MeV/u. The beams are then injected into the PS (5.9 GeV/u) and follow the same accelerator chain as the protons into the SPS (177 GeV/u) and LHC, where they collide at a centre-of-mass energy of 2.76 TeV/u.

3.1.2 Machine Design and Current Status

The LHC is a particle-particle collider. Unlike particle-antiparticle colliders in which the two beams can share the same physical space, the LHC needs two separate rings for the protons to circulate. Due to space restrictions in the tunnel, the two beam channels and two sets of coils, producing opposite magnetic fields, are contained within a single mechanical structure of a twin-bore super-conducting magnet, as illustrated in figure 3.2. The entire system is cooled down to below 2 K by means of superfluid helium and placed under a vacuum. This allows the magnets to reach a field strength above 8 T necessary to bend the particles around the LHC circumference.

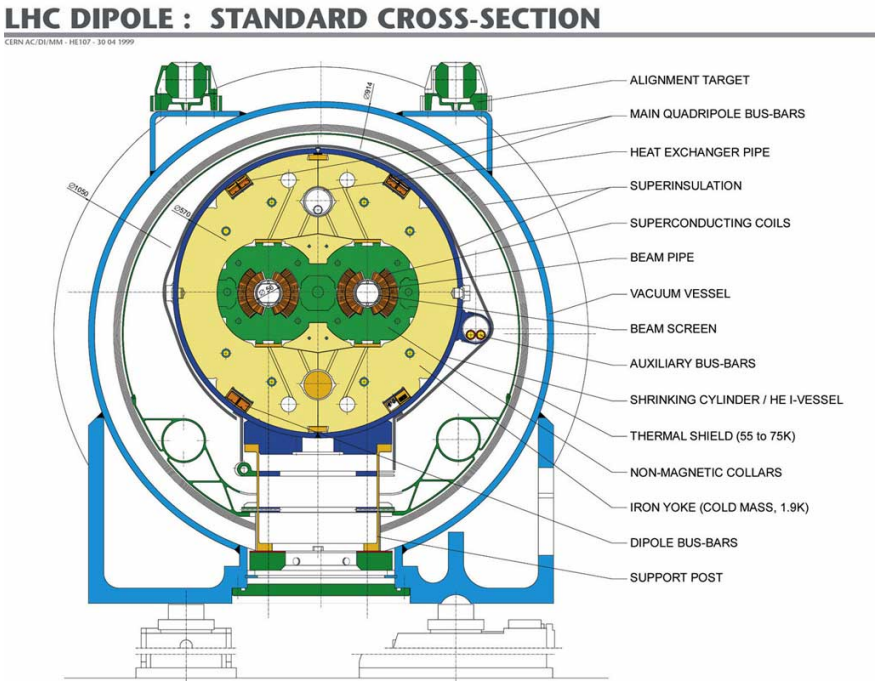


Figure 3.2: Schematic diagram of an LHC dipole magnet. From [47].

The particles are accelerated to the maximum collision energy with Radio Frequency (RF) super-conducting cavities. Once the energy is reached, these cavities are responsible for maintaining the intensity of the beams by keeping the particle bunches close together.

The total number of events per second delivered by the machine depends on its luminosity L and the cross-section σ_{event} of the process under investigation as

$$N_{events} = L \cdot \sigma_{event}. \tag{3.1}$$

The machine instantaneous luminosity is given by

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi \epsilon_n \beta^*} F, \tag{3.2}$$

where N_b corresponds to the number of particles per bunch, n_b to the number of bunches per beam, f_{rev} to the revolution frequency, γ_r to the relativistic factor $(1 - \beta^2)^{-1/2}$, ϵ_n to the beam emittance, β^* to the beta function at the collision point and F is just a reduction factor to account for the beam crossing angle.

The integrated luminosity $\mathcal{L}$, i.e. the total number of collected events over a certain period of time, is obtained by

$$\mathcal{L} = \int L \cdot \mathrm{dt}. \tag{3.3}$$

The LHC began operations in November 2009 with the first 900 GeV proton beams colliding a few days after they first circulated the rings. A world-record energy was achieved before the 2009 winter shutdown with 2.36 TeV centre-of-mass collisions. In March 2010, activities resumed and the energy of the proton collisions was increased to 3.5 TeV *per beam*. Figure 3.3 is an event display of one of the first 7 TeV collisions registered by the ATLAS detector. Since then, the LHC has performed very well, delivering by the end of October of 2011, ahead of schedule, around 5.61 fb^{-1} of proton data to the high-luminosity experiments. The maximum instantaneous luminosity recorded was 3.65×10^{33} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ with 1,854 colliding bunches. Short two-

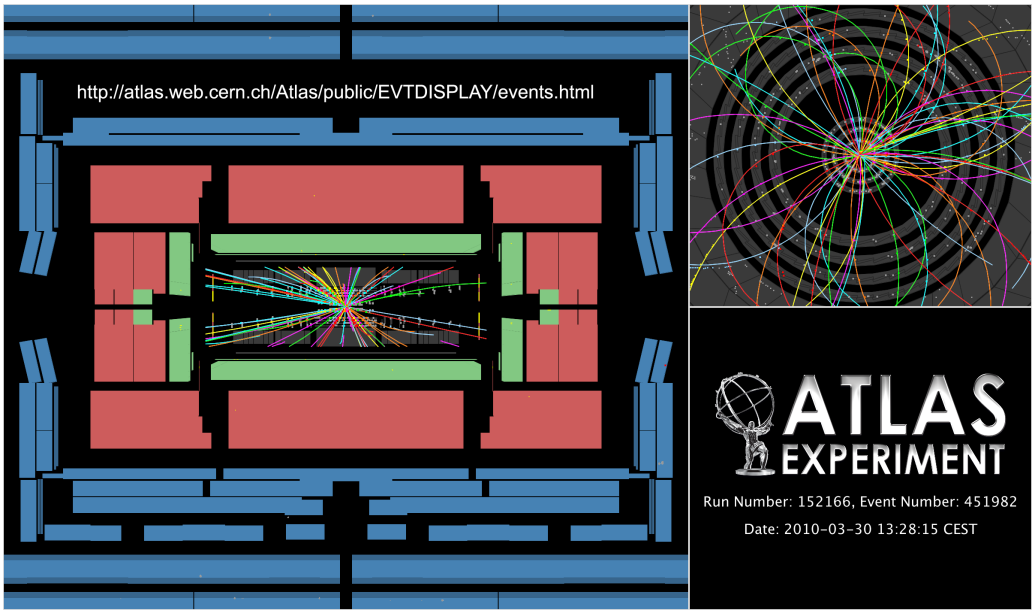


Figure 3.3: Event display of a 7 TeV collision in the ATLAS detector. From [49].

week runs with lead ions took place in November of 2010 and 2011, before the yearly winter shutdowns. The current plan is to continue running at 7 TeV and then increase the energy to 8 TeV until 2013 (with a few weeks of heavy ion collisions in-between) when there will be a long shutdown in preparation for proton collisions at the design energy of 14 TeV.

3.1.3 The LHC Experiments

There are 6 experiments at the LHC, each with unique detector technologies designed to study different aspects of high-energy particle collisions.

High-luminosity Experiments

There are two high-luminosity experiments at the LHC: ATLAS and CMS. The aim is to record a peak luminosity of 10^{34} cm^{-2}s^{-1}. The ATLAS detector will be described in detail in section 3.2.

The *Compact Muon Solenoid* (CMS) [50], just like ATLAS, is a general-

purpose detector. It studies a wide range of processes from the production of the Higgs boson to the discovery of new physics. The tracking and calorimetry systems are located inside a very powerful solenoid magnet, providing very good momentum resolution and electron/photon identification, and also making the overall size of the detector "small" compared to its weight (12,500 tonnes). Over 3,500 scientists work for CMS, making it the largest scientific collaboration in the world.

Low-luminosity Experiments

Operating at a lower luminosity are the LHCb, TOTEM and LHCf experiments.

The *Large Hadron Collider beauty experiment* (LHCb) [51] is a particle detector designed specifically to study the decays of B-mesons (particles which contain the b quark) to try to shed some light on the asymmetry between matter and antimatter. It aims to register peak luminosities of $2{\times}10^{32}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Despite being a collider detector, its architecture resembles that of a fixed-target experiment focusing on the study of very forward particles. To distinguish the B-mesons from the rest of the produced particles, LHCb has a vertex detector located only 8 mm away from the interaction point. It also contains Ring Imaging Cherenkov counters (RICH) for the correct identification of the decay products of the B-mesons.

TOTEM stands for *TOTal cross-section, Elastic scattering and diffraction dissociation Measurement at the LHC* [52]. It is a small experiment dedicated to the measurement of the proton-proton total cross-section and to probing the proton structure in the very forward regions. Roman Pot stations are located between ± 147 and ± 220 m from the CMS interaction point and can localise the trajectory of protons within 1 mm from the beam-pipe. It will

operate at a peak luminosity of 10^{29} cm^{-2}s^{-1}. Its results complement the rest of the LHC experiments, particularly those from CMS.

The smallest LHC experiment is the *Large Hadron Collider forward* (LHCf) [53] experiment. Neutral particles in the forward direction, produced in the collisions at the centre of ATLAS, are detected by two sets of sampling calorimeters at ±140 m from the interaction point. The goal of this experiment is to calibrate the models for hadron interactions used in high-energy cosmic-ray simulations. Designed to withstand low luminosities ($< 10^{30}$ cm^{-2}s^{-1}), due to the radiation damage suffered by the sensors, the LHCf detectors collected proton-collision data during a short period of time between 2009 and 2010 at 900 GeV and 7 TeV. On July 2010 the detectors were removed from the LHC tunnel.

Heavy-ion Experiment

The LHC has one dedicated heavy-ion experiment: ALICE, *A Large Ion Collider Experiment* [54]. It is designed to study lead ion collisions in the search for the quark-gluon plasma, a state of matter in which the quarks and gluons inside the nucleus are no longer bound due to the extreme temperatures reached. This is believed to have been the condition of the Universe right after the Big Bang. Aiming to register a peak luminosity of 10^{27} cm^{-2}s^{-1}, the ALICE detector has very precise sub-systems for identifying the flavour composition of the ion collisions (particle identification) and a Time Projection Chamber (TPC) for very efficient track reconstruction.

3.2 ATLAS - A Toroidal LHC ApparatuS

ATLAS [55] is the largest of the two general-purpose detectors at the LHC. It is 44 m long and 25 m high and it weighs approximately 7,000 tonnes. Col-

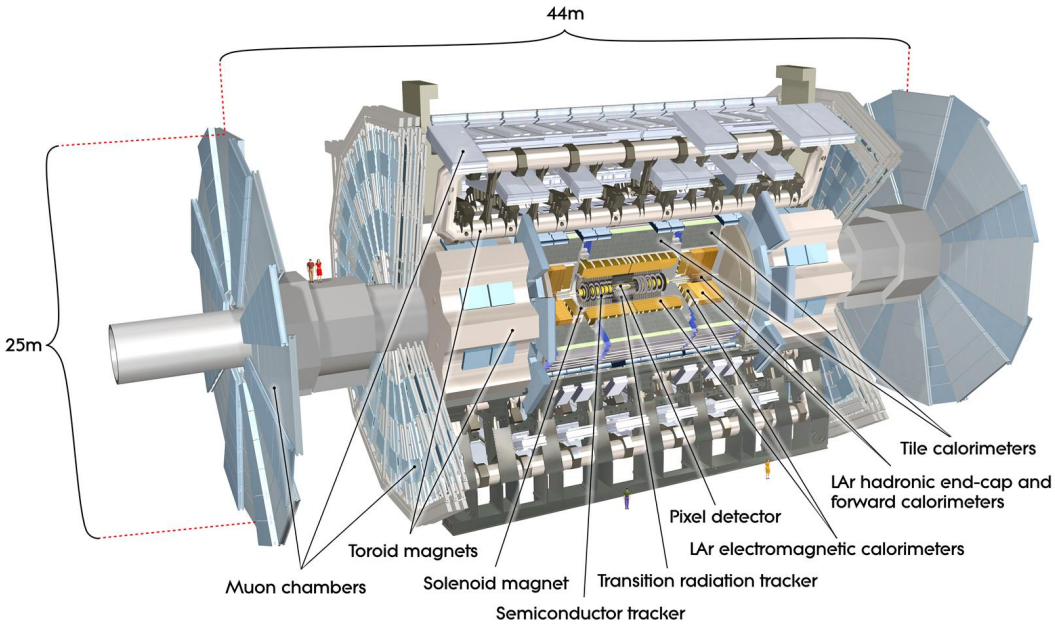


Figure 3.4: Cut-away schematic view of the ATLAS detector. From [55].

lision events are reconstructed by layers of precise tracking, calorimetry and muon systems surrounding the interaction point (figure 3.4). These subsystems will be described in more detail in the following sections with particular emphasis on the tracking detectors since they are the most relevant for the measurements presented in this thesis.

The physics goals of the ATLAS experiment range from stringent tests of the Standard Model and its theories, to the discovery of the Higgs boson and new physics such as SUSY and extra-dimensions. The operation of such a complex detector and the wide variety of processes to analyse results in more than 3,000 scientists being part of the ATLAS Collaboration.

Coordinate System

The ATLAS detector is forward-backward symmetric with respect to the interaction point. The nominal interaction point is defined as the geometrical

centre of the detector and is considered as the origin of the ATLAS Cartesian coordinate system. The positive x-axis is defined as pointing from the interaction point towards the centre of the LHC; the positive y-axis points upwards from the interaction point and the positive z-axis is defined along the beam pipe in the anti-clockwise direction.

3.2.1 Inner Detector

At the heart of ATLAS, closest to the interaction point and covering the region $|\eta| < 2.5$, lies the Inner Detector (ID) responsible for measuring the momentum of charged particles and reconstructing their trajectories (referred to as tracks), and the vertices from which they came. To be able to deal with the high density of particles produced in the collisions at the LHC, the Inner Detector is designed to be fast and radiation-hard and to provide robust pattern recognition and good spatial resolution.

The ID is immersed in a 2 T magnetic field from a super-conducting solenoid that bends the trajectories of the charged-particles and allows the determination of the momentum. It consists of three sub-systems: a silicon pixel detector (Pixel), a silicon micro-strip detector (SCT) and a transition radiation tracker (TRT). A schematic view of the ID is shown in figure 3.5 and a plane view of a quarter-section with dimensions and η coverage is shown in figure 3.6.

Pixel Detector

The Pixel sub-system is the closest to the beam pipe. Thanks to its high granularity it is capable of resolving primary and secondary vertices.

The Pixel modules are arranged in three concentric cylinders in the barrel and three discs perpendicular to the beam axis at each of the end-cap regions

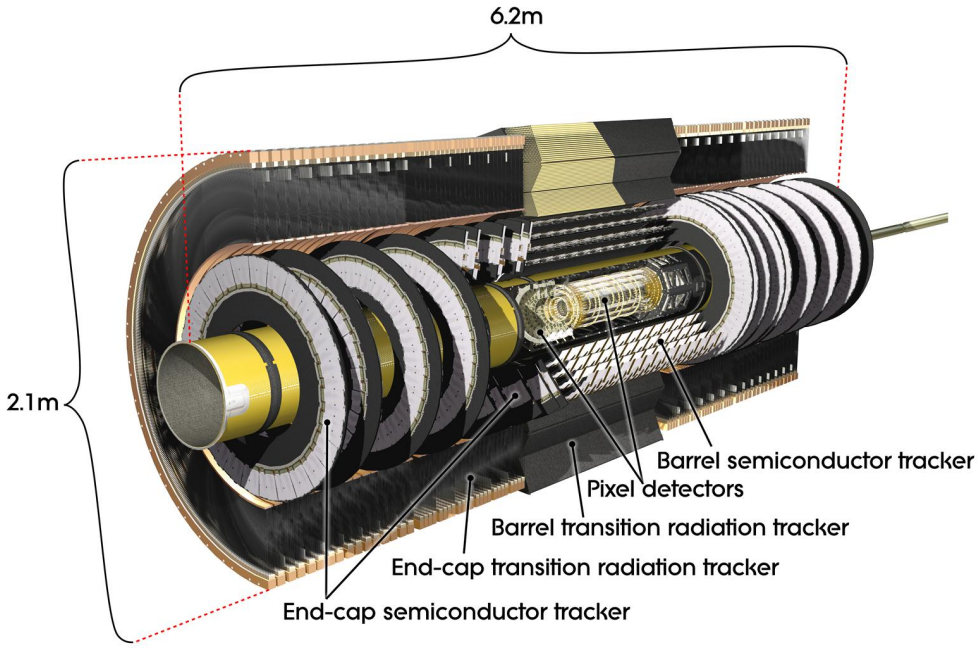


Figure 3.5: Cut-away schematic view of the ATLAS Inner Detector. From [55].

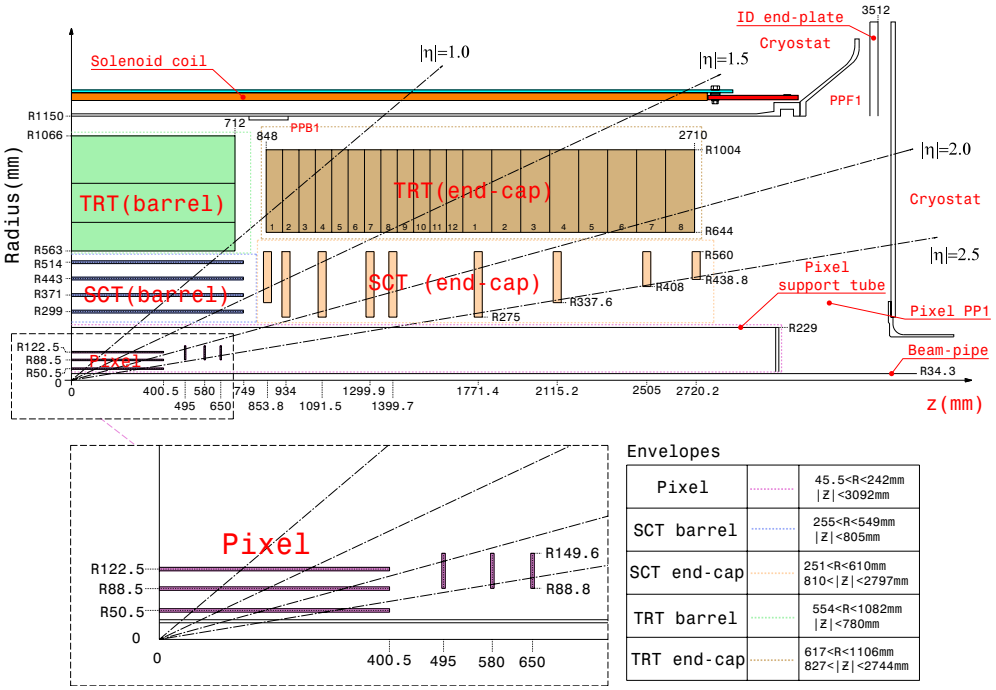


Figure 3.6: Layout of a quarter-section of the Inner Detector and its components. From [55].

of the detector. Typically, three Pixel layers are crossed by a charged particle as it travels through ATLAS. The innermost barrel layer, the *b*-layer, located at 50.5 mm from the beam line, is the most exposed to radiation and as such will have to be replaced after approximately three years of operation at design luminosity.

A Pixel module consists of a silicon sensor 250 μm thick, segmented into small rectangles, or *pixels*, in the $R-\phi$ and z directions. When a charged particle goes through the sensor, it liberates electron-hole pairs that drift under the electric potential (initially 150 V) towards the pixels. This signal is read and, if it is above the threshold of 4,000 e^-, its information stored for later use in the track reconstruction. The Pixel sensors have intrinsic accuracies of 10 μm ($R-\phi$) and 115 μm (z) in the barrel layers and 10 μm ($R-\phi$) and 115 μm (R) in the end-cap discs. To reduce noise, the silicon sensors in both the Pixel and SCT detectors are cooled down to between $-5°$ C and $-10°$ C.

Semi-Conductor Tracker (SCT)

Surrounding the Pixel detector, the Semi-Conductor Tracker contributes to the momentum and vertex position measurements and to the calculation of the impact parameters (see section 4.2.1). It also has a high granularity, necessary for the pattern recognition in the track reconstruction.

The SCT is divided in four barrel layers and nine discs in each of its two end-cap regions. An SCT module consists of two pairs of single-sided p-in-n silicon microstrip sensors glued back-to-back with a 40 mrad stereo rotation angle. The sensors are 285 μm thick and operate at an initial voltage of 150 V. The SCT opted for a binary readout, which means that a strip will only "fire" if it collects enough charge above the 1 fC threshold. This threshold

value is chosen to maximise the efficiency and minimise the noise (the noise occupancy is required to be $< 5 \times 10^{-4}$). This has an effect on the intrinsic measurement accuracies: 17 μm in the $R\phi$-plane and 580 μm in the z-plane (R-plane) for barrel (end-cap) modules.

In total, there are 2,112 barrel and 1,976 end-cap modules, the latter of different shapes depending on their location within the discs. Each side of a module has 768 strips and in the case of the rectangular barrel sensors, they are separated by an 80 μm pitch and run parallel to the beam axis, while for the end-cap modules the strips run radially. On average, eight strip layers are crossed by each track which is equivalent to four space-points (intersection of two strips from opposite sides of the module).

Transition Radiation Tracker (TRT)

The TRT is the outermost system of the Inner Detector. It comprises layers of straw tubes filled with a gas mixture of xenon, carbon dioxide and oxygen: $XeCO_2O_2$. With an average of 36 hits per track, the TRT provides continuous tracking to enhance the pattern recognition and the momentum resolution. Electron identification complementary to that provided by the electromagnetic calorimeter, up to $|\eta| < 2.0$, is possible due to the transition radiation material planes interleaved between the straws.

The straw aluminium coated tubes have a diameter of 4 mm. Within 300 μm from the centre of the straws, a tungsten wire with a diameter of 30 μm collects the charge clusters released by an incoming particle. In the barrel the straws are 144 cm long and are parallel to the beam direction, while in the end-caps the average length is 37 cm and the straws are arranged radially in wheels. The intrinsic accuracy in the $R - \phi$ direction is 130 μm.

The TRT is divided into three barrel layers, each containing 32 modules,

and two sets of wheels in each end-cap. A barrel module is an array of straws surrounded by polypropylene fibres acting as transition radiation material. An end-cap wheel contains 8 layers of straw tubes, each integrated by 768 straws, followed by 15 μm polypropylene radiator foils.

3.2.2 Calorimetry

The ATLAS calorimeters, illustrated in figure 3.7, are responsible for measuring the energy of neutral and charged particles by alternating active and passive materials. The choice of detector technology depends on whether the particles interact via the electromagnetic or strong forces; this is the reason why in ATLAS there are separate electromagnetic and hadronic calorimeters that, together, cover the range $|\eta| < 4.9$.

The calorimeters are designed so that showers are contained inside them (hadronic showers are longer and wider than electromagnetic ones) to accurately measure the energy and to limit the flux into the muon spectrometer. The calorimeters have a very important role to play in the search for new phenomena due to their jet energy and E_T^{miss} measurements.

Electromagnetic Calorimeter

The electromagnetic calorimeter (EM) uses lead plates as an absorber and liquid argon (LAr) as the active material. It has high granularity and hence provides excellent energy and position resolution in the range $|\eta| < 3.2$. The electrodes have an accordion shape which provides complete symmetry in the ϕ direction without any cracks.

It is divided in two half-barrel regions ($|\eta| < 1.475$) separated by a 4mm gap at $z = 0$, and two coaxial wheels, inner ($1.375 < |\eta| < 2.5$) and outer ($2.5 < |\eta| < 3.2$), in each end-cap side. In the region where the EM calorimeter

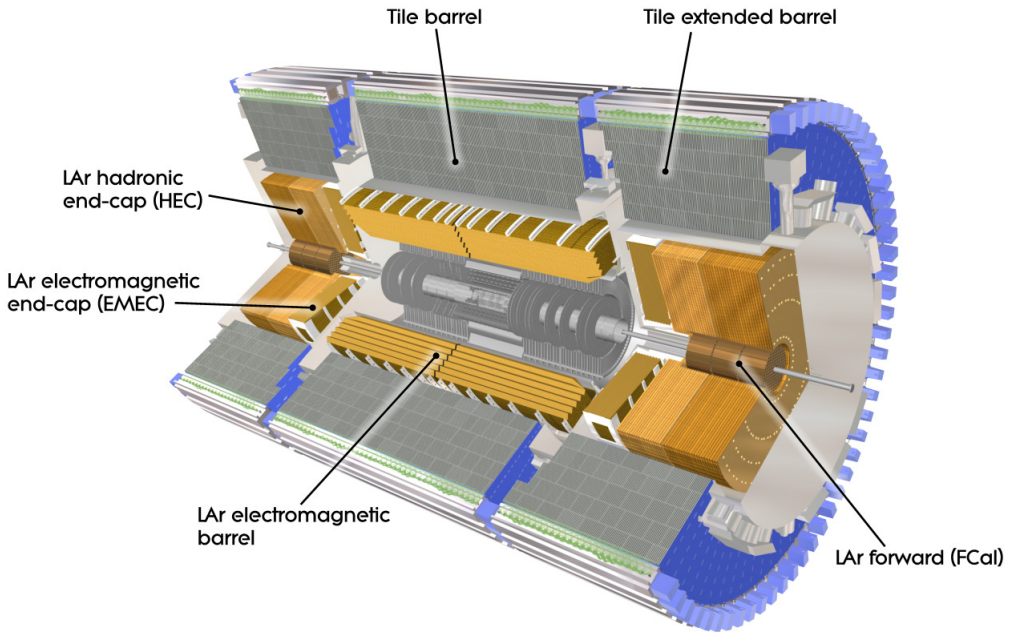


Figure 3.7: Cut-away schematic view of the ATLAS Calorimeters. From [55].

overlaps with the ID ($|\eta| < 2.5$), the innermost compartment, or *sampling*, is finely segmented to allow for precise identification of photons and electrons.

Hadronic Calorimeters

Two different technologies are used for the hadronic calorimeters.

In the barrel, directly outside the EM, the Tile Calorimeter operates. The active materials are scintillating tiles and the absorbers are steel plates. It is segmented into one central barrel with an η coverage of $|\eta| < 1.0$ and two smaller extended barrels on either side covering the $0.8 < |\eta| < 1.7$ region. The tiles are oriented radially, perpendicular to the beam direction.

The Hadronic End-cap Calorimeter (HEC), as the name suggests, is located in the end-caps. LAr is again used as an active material but the interleaved absorber plates are made of copper. It consists of two wheels on either side and it extends from 1.5 to 3.2 in $|\eta|$, overlapping with both the Tile and

Forward Calorimeters, reducing the material budget in the transition region between the barrel and the end-caps.

Forward Calorimeter

Extending the coverage to the high pseudorapidity region $3.1 < |\eta| < 4.9$ is the Forward Calorimeter (FCal), located inside the same cryostat as the end-cap calorimeters (EM+hadronic). It is split into three compartments. The first one, dedicated to electromagnetic measurements, uses LAr and copper as active and passive materials, respectively. The other two sections are in charge of measuring hadronic showers and use tungsten as an absorber with LAr filling the space between the rods.

3.2.3 Muon Spectrometer

The Muon Spectrometer (MS) is designed to both trigger and measure the momentum of muons in the range 0.03 - 3 TeV. High transverse momentum muons are a signature of new and interesting physics, as such, the muon system is a fundamental element of the ATLAS detector.

The MS is integrated by four different types of detection chambers and a magnet system (figure 3.8). For tracking and momentum measurements in the bending (yz-) plane there are the Monitored Drift Tube chambers (MDTs) and the Cathode-Strip Chambers (CSCs). Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) are used for bunch-crossing identification and measurements of track coordinates in the non-bending (xy-) plane.

The magnetic field necessary to bend the trajectories of the muons is provided by three (1 barrel and 2 end-cap) super-conducting air-core toroid magnets. In the pseudorapidity range $0 < |\eta| < 1.4$, the barrel toroid provides 1.5 to 5.5 T·m of bending power, while in the $1.6 < |\eta| < 2.7$ range, the bending

power provided by the end-cap toroids varies between 1 and 7.5 T·m. In the transition region, a combination of the two magnetic fields is used to bend the tracks.

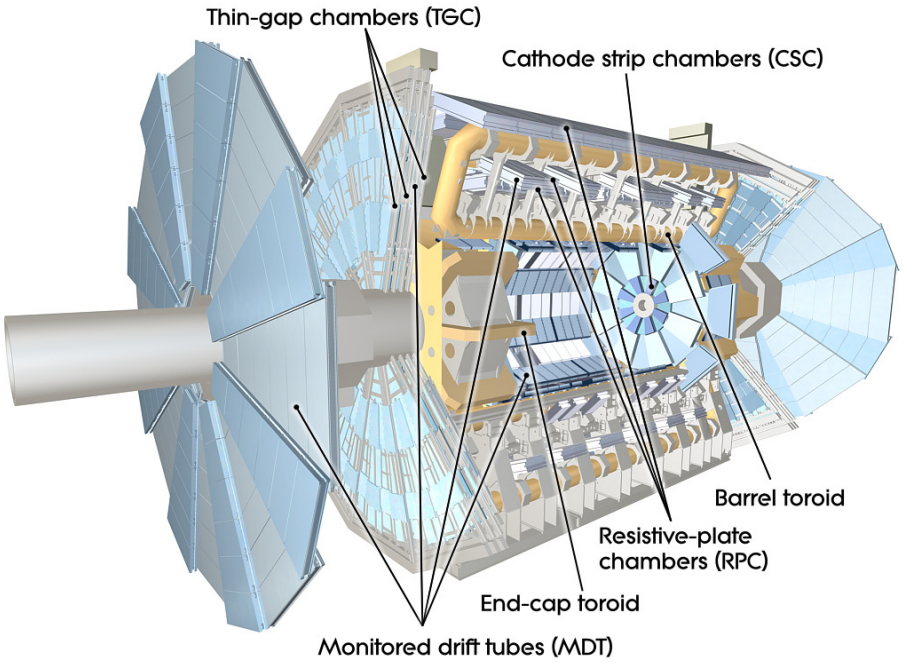


Figure 3.8: Cut-away schematic view of the ATLAS Muon Spectrometer. From [55].

The MDTs are arranged in three layers that cover the range $|\eta| < 2.7$. In the innermost layer, between 2.0 and 2.7 in pseudorapidity, the MDTs are replaced by CSC multi-wire proportional chambers to manage the high rates and backgrounds in this forward region. The trigger system covers the range $|\eta| < 2.4$ and has a timing resolution between 1.5 and 4 ns. RPCs are used in the barrel and TGCs in the end-caps of the detector.

3.2.4 Forward Detectors

The very forward region of ATLAS is covered by three small detectors. LUCID (LUminosity measurement using Cerenkov Integrating Detector) and ALFA

(Absolute Luminosity For ATLAS) determine the luminosity delivered to the experiment at ±17 and ±240 m from the interaction point, respectively.

The third system is the Zero Degree Calorimeter (ZDC), located at ±140 m from the centre of ATLAS, and it evaluates the centrality of the heavy-ion collisions by measuring neutral particles at $|\eta| \geq 8.2$ with alternating layers of quartz rods and tungsten plates.

3.2.5 Trigger and Data Acquisition

Every 25 ns[1], protons collide at the centre of ATLAS. The amount of data produced is enormous and as the luminosity increases, this will only get more complicated due to the occurrence of more than one proton interaction per bunch crossing. The trigger and data acquisition systems (DAQ) are designed to select only interesting events among the initial 40 MHz rate, to optimise the storage and analysis capabilities of the experiment.

The ATLAS trigger has three levels: Level-1 (L1), Level-2 (L2) and the Event Filter (EF). Together, L2 and EF form the High-Level Trigger (HLT). The L1 trigger searches for signatures from high-p_T muons, jets, electrons, photons, missing energy, among others, using only a subset of the ATLAS detector (Muon Spectrometer and calorimeters) to define a Region of Interest (RoI). At L1 the maximum rate that can be managed lies between 75 and 100 kHz, and the trigger decision must reach the front-end electronics 2.5 μs after the interaction.

The L2 trigger is seeded by the RoIs and uses detailed information from the detector elements within them. The rate is reduced to 3.5 kHz with an average event processing time of 40 ms. Finally, the EF uses offline analysis

[1]During the 2010-2011 data-taking period, the LHC operated with a 50 ns clock.

algorithms to further reduce the rate to approximately 200 Hz and, on average, processes events in 4 seconds. The events that pass the EF selection criteria are used for offline analysis. The HLT uses the full granularity of the muon spectrometer and calorimeters and also information from the Inner Detector, which results in better threshold cuts and particle identification.

The DAQ is in charge of buffering the data at the L1 rate (readout system) and transferring it to the L2. For events going through to the last stage, the DAQ performs event-building so that they can be analysed by the EF and, if selected, stores them permanently. Furthermore, the DAQ is responsible for sending configuration, controlling and monitoring the detector during data-taking.

The interface to monitor the status of the hardware is called the Detector Control System (DCS). Operational parameters like voltage, cooling, temperature and safety can be accessed by the operator and, if necessary, allow for intervention. In addition, the DCS and the DAQ communicate to each other to synchronise the detector status with data-taking.

3.2.6 Current Status

The ATLAS detector has had a very good performance with nearly all of its channels fully operational (figure 3.9). The total integrated luminosity delivered by the LHC in 2011, as of November, is approximately 5.61 fb^{-1}, of which 5.25 fb^{-1} was recorded by the experiment. This translates into an overall data taking efficiency of 93.5% (figure 3.10).

Subdetector	Number of Channels	Approximate Operational Fraction
Pixels	80 M	96.4%
SCT Silicon Strips	6.3 M	99.2%
TRT Transition Radiation Tracker	350 k	97.5%
LAr EM Calorimeter	170 k	99.8%
Tile calorimeter	9800	96.2%
Hadronic endcap LAr calorimeter	5600	99.6%
Forward LAr calorimeter	3500	99.8%
LVL1 Calo trigger	7160	99.9%
LVL1 Muon RPC trigger	370 k100	99.0%
LVL1 Muon TGC trigger	320 k	100%
MDT Muon Drift Tubes	350 k	99.7%
CSC Cathode Strip Chambers	31 k	97.7%
RPC Barrel Muon Chambers	370 k	97.0%
TGC Endcap Muon Chambers	320 k	97.9%

Figure 3.9: Operational status of the ATLAS detector as reported in [56]. The total number of channels and the approximate operational fraction are shown for each sub-system.

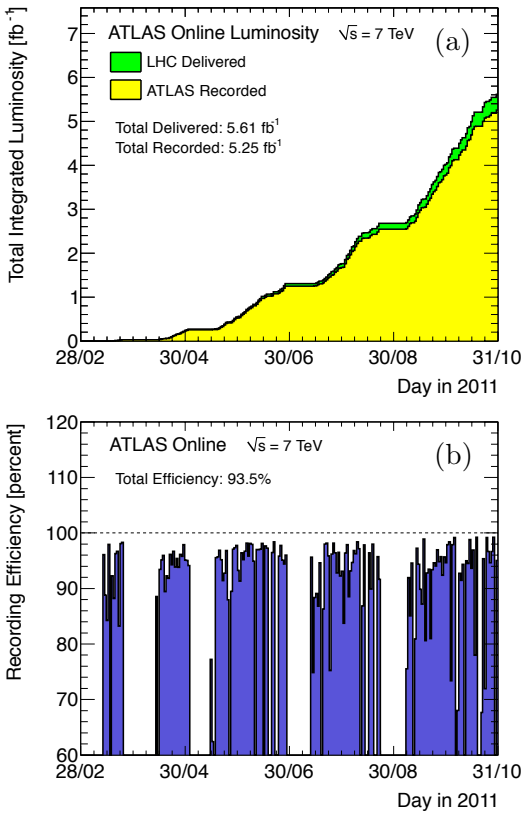


Figure 3.10: (a) Total integrated luminosity delivered by the LHC in 2011 and (b) data-taking efficiency of the ATLAS detector as reported in [57].

Chapter 4

Data Reconstruction

4.1 The ATLAS Offline Software

The ATLAS software framework, ATHENA [58], is adapted from the GAUDI [59] framework originally developed by the LHCb collaboration. It originates from a component-based *blackboard* architecture in which the input data, services and tools are located in a common in-memory database and loaded as libraries at the job configuration level by a Python script (referred to as *job options* file). After initialisation of an ATHENA job, the algorithms, coded in C++, are executed in sequential order on an event-by-event basis. Finally, algorithms are terminated, objects are deleted and the output is written and persistified. The modular structure of ATHENA allows for the development of tools and packages, separated into different projects, that can be constantly updated and frozen into new versions, or releases, of the software.

4.1.1 Detector Geometry Description

For an accurate reconstruction of either real or simulated data, a precise geometrical description of the ATLAS detector is needed; it relies on two

software components: a relational database containing the configuration and properties of the detector, and a set of geometrical primitives. A tag with the detector geometry version is specified either at simulation time or at data recording time. This tag is persistified along with the data and provides access to the configuration parameters of the detector via the relational database.

The library of geometrical primitives is provided by the GeoModel toolkit [60]. This software is designed to describe large and complex detectors while optimising the memory consumption. In the ATLAS GeoModel implementation [61], the material and readout geometry of each sub-detector is modelled independently.

4.1.2 Simulation Infrastructure

A very important part of the ATLAS programme is the evaluation of the detector response to different physics processes and the comparison of the experimental results to the available models. For this, ATHENA provides the framework for Monte Carlo (MC) production which is carried out in three main steps [62]: generation, simulation and digitisation. The subsequent triggering and reconstruction are performed using the same algorithms used for real ATLAS data (figure 4.1).

Generation

Several event generators are supported in ATHENA, each with different approaches to the modelling of hard processes, initial and final state radiation, multi-parton interactions, hadronisation, etc. The output of these generators, the event record, is converted into the HepMC format [63] which stores the complete event history as the MC Truth information.

At this stage, no geometry description of the detector is necessary since all

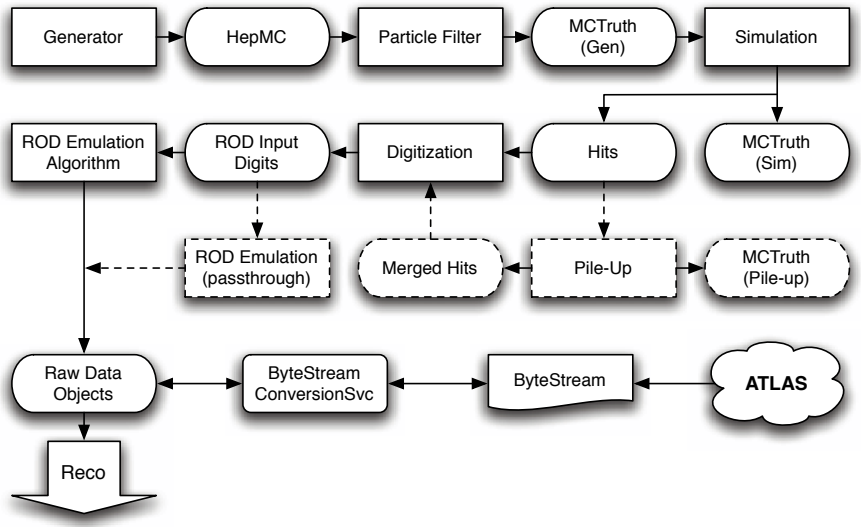


Figure 4.1: Diagrammatic representation of the simulation process in the ATLAS software. Once RDOs are produced, the same reconstruction algorithms can be applied to both real and Monte Carlo simulated data. From [58].

stable particles, defined by a proper lifetime $c\tau > 10$ mm, are handed over to the simulation. Unstable particles are decayed by the generator and the effect of the magnetic field and material is neglected. One additional feature that can be done at the event generation level is filtering, in which only particles that satisfy certain properties (e.g. decay channel, energy, momentum) are propagated into the simulation chain.

Simulation

The simulation of the detector response and physics processes occurring inside the detector material is carried out via the Geant4 particle simulation toolkit [64, 65]. This is the most CPU time consuming stage of the production process. Only particles that are within a specified $\eta - \phi$ range are propagated through the geometry; very forward particles are not simulated because they increase the simulation time by, approximately, a factor of 2.

The MC Truth record produced at event generation is copied and extended

with information from simulation. The final output of the simulation stage is a hit file which contains metadata (describing the configuration), the extended MC Truth and a collection of hits for each sub-detector.

Digitisation

In this last step, hits produced by the simulation are converted into detector responses called *digits*, which are the actual output format of the ATLAS detector. These digits are fed to simulated Read-Out Drivers (ROD) that emulate the front-end electronics in the detector cavern. The output is written as Raw Data Objects (RDOs) over which the same trigger and reconstruction algorithms of data can be run. As an additional feature, these RDOs, unlike the ones from real data, contain all the MC Truth information.

At the LHC, more than one proton-proton interaction per bunch crossing can occur. The digitisation framework has the machinery to add these extra interactions by reading different types of events and overlaying hits from each of them before building the digits.

Fast Simulation

The complexity and level of detail of the Geant4 full simulation makes it difficult to produce large amounts of Monte Carlo data for all analyses. Currently there are three approaches to fast simulation in ATLAS, all of which have been validated against the full simulation chain:

- **FastG4** [66, 67] - removes low energy electromagnetic particles and replaces them with pre-simulated showers.
- **ATLFAST-I** [68, 69] - truth objects are smeared by detector resolution effects; the speed of simulation is increased by a factor of 1,000.

- **ATLFAST-II** [68, 69] - the goal is to simulate events as fast as possible while still being able to use the normal ATLAS reconstruction algorithms. It uses Fatras (Fast ATLAS Tracking Simulation) [70] for Inner Detector and Muon Spectrometer tracks and FastCaloSim (Fast Calorimeter Simulation) for the calorimeter responses.

4.1.3 Reconstruction and Analysis Preparation

From the RDOs, one can derive parameters that can be used for physics analysis using algorithms for tracking (section 4.2), calorimeter towers, etc. This is called reconstruction and in ATLAS it consists of different stages that produce different output datasets. The first and most general type of data are the ESDs, or Event Summary Data, and they contain detailed information to allow particle identification, track re-fitting, jet calibration, etc. The AODs, Analysis Object Data, are the second stage and are a summary of the ESDs and contain sufficient information for most analyses. Finally, there are the D3PDs, produced from AODs, that have a flat ntuple structure and are produced to suit the requirements of the different physics groups. Most analyses are carried out on D3PDs.

4.2 Track Reconstruction

The efficient reconstruction of the trajectories of charged-particles is fundamental to all particle physics analyses, and in particular to the measurement presented in this thesis. As discussed in section 3.2, there are two tracking systems in ATLAS: the Inner Detector (ID) and the Muon Spectrometer (MS). Despite the different conditions these elements are subjected to, the track reconstruction software that has been developed in ATLAS uses a common

Event Data Model (EDM) [71] and a consistent geometrical description of the detector [72]. It is called New Tracking, or simply, NEWT [73].

One of the main features of NEWT is its modular structure. The complex process of track reconstruction is factorised into modules, each performing specific and independent tasks. This results in a very flexible software that can easily be maintained and extended.

This section focuses on the sequences for track reconstruction in the ID. First, a brief description of the ATLAS Tracking EDM and the variables used for track parameterisation is presented.

4.2.1 EDM and Track Parameterisation

The main feature of the tracking EDM is a common track object that describes both ID and MS tracks and, in addition, can be used in online processing and offline reconstruction. This track object can contain multiple parameters describing, for example, the position of the track at different surfaces, the trajectory it follows, which algorithm was used to create it, the quality of the track fit, and many others.

In the ATLAS tracking EDM parameterisation, tracks can be represented locally by a set of five parameters

$$\tau_i = (l_1, l_2, \phi, \theta, q/p), \tag{4.1}$$

where (l_1, l_2) denote two coordinates in the intrinsic frame of the surface where the measurement is done, and $(\phi, \theta, q/p)$ represent the momentum of the track in the global frame. This representation differs slightly from the $(\phi, \cot\theta, p_{\mathrm{T}})$ one, commonly used for helix parameterisations in particle physics experiments, simply because the local parameterisation cannot be restricted to a

specific detector; the transition between the inner tracker and the muon systems is affected by a non-homogeneous magnetic field, which means that p_{T} is not a constant of motion.

Of all the different representations available in the EDM, the perigee helix parameterisation is of particular importance. The perigee represents the point of closest approach of the track to a specific point, that can be the nominal interaction point or the position of the primary vertex, and it can be described by the following parameters [74], illustrated in figure 4.2:

- d_0 - transverse impact parameter; distance of closest approach in the transverse (xy-) plane.
- z_0 - longitudinal impact parameter; z coordinate of the point of closest approach.
- ϕ - azimuthal angle of the momentum of the track at the perigee; measured in the xy-plane.
- θ - polar angle of the track; measured in the zd-plane, where $d = \rho\phi$ with ρ the radius of curvature of the helix.
- q/p - charge of the track divided by its momentum.

4.2.2 Inside-Out Tracking Sequence

The primary ID track reconstruction sequence starts from the silicon trackers and moves outwards into the TRT. It exploits the high granularity of the Pixel and SCT detectors to identify tracks close to the interaction point.

In the first step, raw information from the silicon detectors is converted into three-dimensional representations called space-points. In the case of the Pixel detector, a space-point is defined by a single hit in the sensors. For

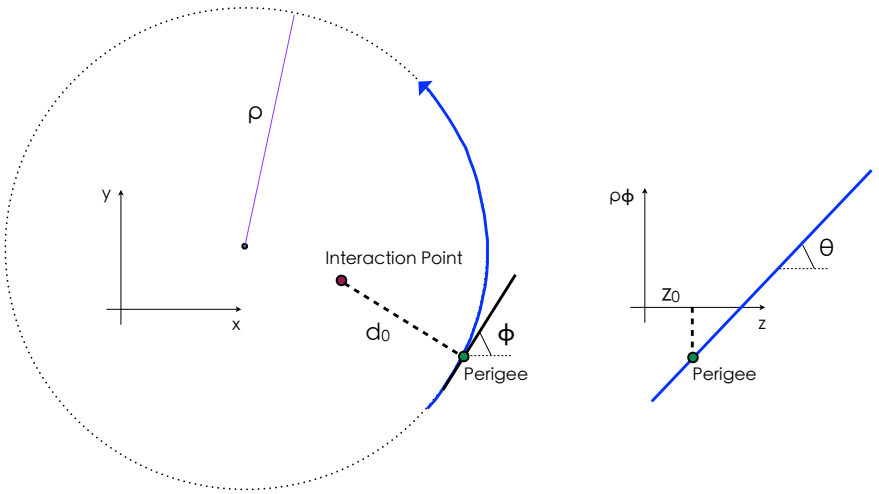


Figure 4.2: Track parameters at the perigee.

the SCT it is required to have a hit on opposite sides of a module so that, thanks to the stereo rotated sensors, two dimensional measurements can be constructed.

Next, with the space-points located in any of the three Pixel layers and the first SCT layer, track seeds are formed from which paths are built. Any silicon cluster that falls within this path is added to the track candidate by a Kalman filter-smoother formalism [75]. This process results in a large number of track candidates that have to be resolved, before extending them into the TRT, by applying quality cuts and assigning track *scores*. Points are given if the quality of the track fit (χ^2/N_{dof}) is good and if there are many hits associated to it; penalties are applied if, for example, there are holes (no hits where expected) in the trajectory of the track. Candidates with a high score are extended into the TRT and fitted again (several fitting techniques are available; the default is GlobalChi2Fitter [76]), this time using the information from the three sub-detectors, while the ones that fail to pass the quality cuts are neglected from further processing. The score of the track is compared to

the score of the silicon track candidate before extension and the better scoring alternative is kept. These two stages of the track reconstruction sequence are illustrated in figure 4.3.

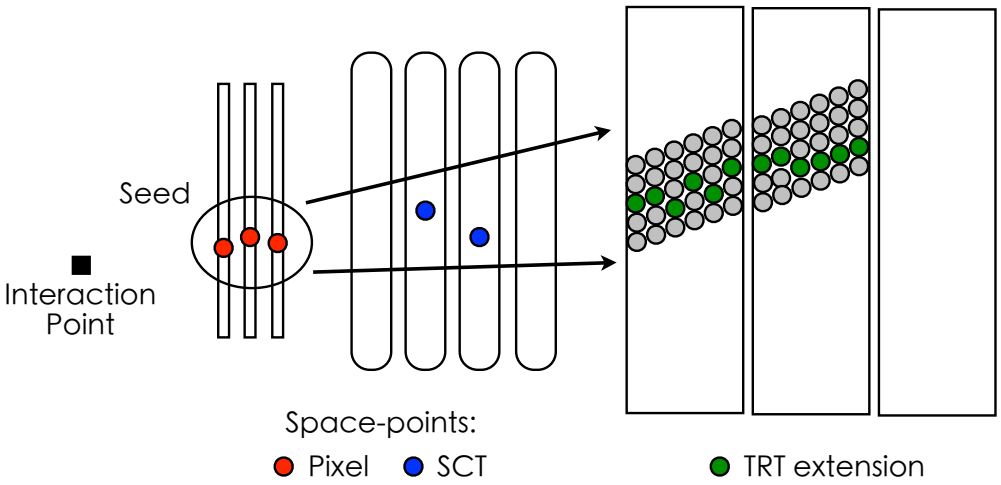


Figure 4.3: Inside-out track reconstruction sequence.

4.2.3 Outside-In Tracking Sequence

The inside-out sequence relies on the identification of seeds to do the pattern recognition. However, it could be the case that, due to ambiguous hits, the seeds are not able to reach the score necessary to go through to the next steps and the track is lost. Furthermore, particles can decay inside the ID volume and hence the tracks associated to these secondary particles will not have enough hits to build a seed in the silicon. To address these issues, a second pattern recognition sequence is available in NEWT which starts from the TRT and *back-tracks* into the silicon detectors.

With the information provided by the drift tubes in the TRT, projection planes are used to do the pattern recognition since no information about the coordinate along the straw direction is available (i.e. no space-point formation). The $R\phi$- and Rz- planes are used for the barrel and end-caps, respec-

tively. Tracks with a $p_{\mathrm{T}} > 500$ MeV will appear as straight lines in these planes. A histogram method, known as the Hough transform [77] is used to find straight line patterns and define track segment candidates, which are later re-processed using a Kalman filter-smoothing formalism. Only segments that have not been already grouped into tracks by the inside-out algorithm are extended into the SCT and Pixel detectors and associated to space-point seeds. Segments that are not associated to any silicon hit are called TRT standalone tracks.

4.2.4 Low-p_{T} Tracking Sequence

The pattern recognition sequences previously described are designed for tracks with a minimum p_{T} of 500 MeV. In order to be able to reconstruct tracks with a lower transverse momentum, the low-p_{T} sequence is implemented in NEWT. It follows the same strategy as the inside-out sequence but, since low-p_{T} tracks bend more under the magnetic field and may not be able to reach the outermost elements of the Inner Detector, the reconstruction requirements are not so strict. This sequence uses only hits that have not been associated to tracks by the previous two algorithms. The minimum p_{T} that can be reconstructed is 100 MeV.

4.3 Vertex Reconstruction

The correct identification and reconstruction of primary vertices in an event is necessary for high-precision measurements. Inside ATHENA, the vertex reconstruction algorithms are executed as a post-processing stage of the tracking sequences described in the previous section.

As in the track reconstruction case, the vertex reconstruction uses a com-

mon object-oriented EDM that contains a set of interfaces and core classes to store the basic information about a vertex object: position in the global detector frame, quality and parameters of the vertex fit, track-to-vertex associations, etc [78].

Two vertex reconstruction strategies are implemented in the ATLAS software: *fitting-after-finding* and *finding-through-fitting*. An example of the former is the InDetPriVxFinder algorithm [78]. It begins by pre-selecting tracks that originate from the bunch-crossing region and ordering them according to their longitudinal impact parameter. Clusters of tracks are stored as primary vertex candidates and reconstructed using a vertex fitter that runs an iterative cleaning procedure to remove outlier tracks ($\chi^2 < 8\%$). Once a track is rejected, it is not used again which means that the number of vertices is entirely determined at the cluster seeding stage.

The second strategy, *finding-through-fitting*, provides a better handling of outlier tracks and is the default in ATLAS. The InDetAdaptiveMultiPriVxFinder algorithm [78] also starts the reconstruction process with a selection of tracks from the interaction region but with them it creates a single vertex seed. The vertex is reconstructed using the AdaptiveMultiVertexFitter [78] and tracks that fail to be added to the vertex are used to create a second seed. The two vertices are re-fitted simultaneously and the process is repeated, with the number of vertices growing after each iteration. Tracks that fail to be associated to any vertex with a $\chi^2 > 1\%$ do not influence the result of the fits.

The EDM provides several vertex fitters as standalone tools that can be used for both primary and secondary vertex reconstruction. A detailed description of these can be found in [78].

Chapter 5

Depletion Depth Studies for the ATLAS SCT

5.1 Charge Depletion in Silicon Sensors

The microstrip sensors of the ATLAS SCT (section 3.2.1) use *p-in-n* silicon technology. A p-type material is obtained by doping the silicon with atoms from the III group of the periodic table. Elements in the III group have three electrons in their valence shell and can attach an electron from the silicon atoms with relative ease (these elements are called *acceptors*). On the other hand, n-type materials are produced by replacing some silicon atoms with atoms from the V group, which characterise for being good *donors* since to lose one of their five valence quarks they require a small amount of energy. The majority carriers in the p- and n-type semi-conductors are holes and electrons, respectively.

When p-type and n-type silicon come together they create a junction [79]. Free carriers at either side of the junction start to diffuse into the opposite region; holes diffuse into the n-side, leaving a net negative charge on the p-side of the junction, and electrons diffuse into the p-side, leaving the n-side of

the junction with a net positive charge. This migration and recombination of electrons and holes leaves a space charge or depletion region that has no free carriers and a small built-in potential, as illustrated in figure 5.1.

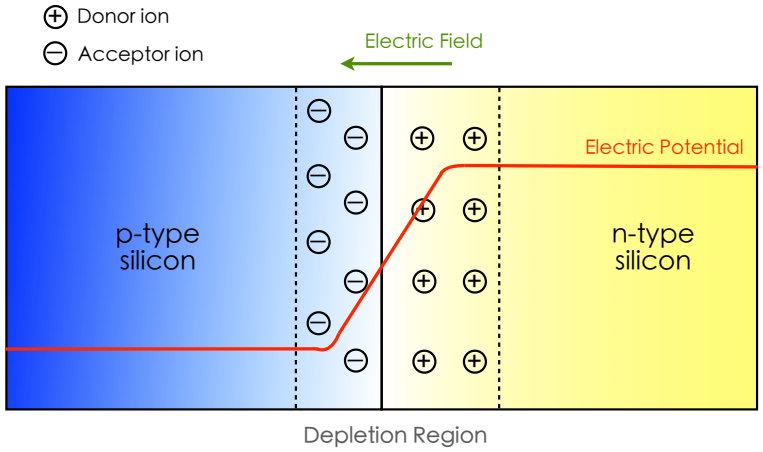


Figure 5.1: Schematic representation of a p-n junction. The immobile acceptor and donor ions in the depletion region generate an electric field and, hence, a built-in potential. The dark coloured regions correspond to neutral material.

The electron-hole pairs liberated by a charged-particle traversing through the depletion region drift under the potential towards the junction and can be detected. In contrast, charge created in the neutral regions quickly recombines with the free carriers and the signal is lost. For this reason, the depletion layer can be thought of as a particle detector. An external reverse-bias potential can be applied to increase the width of the depletion region and, therefore, the active volume of the detector. In fact, if the concentration of acceptor dopants is greater than the donor concentration (p^+n junction) then the depletion region can be very wide on the n-side and very shallow on the p-side. This is the principle of silicon microstrip detectors, such as those in the SCT, in which the bulk of the sensor is of n-type and the implanted strips have been doped with p-type materials (figure 5.2).

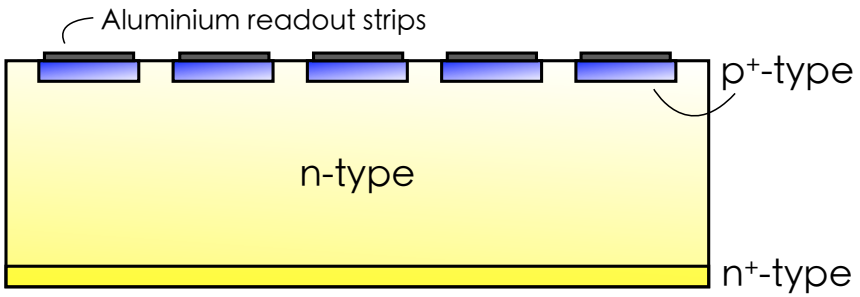


Figure 5.2: Schematic representation of a (p^+, n, n^+) silicon microstrip sensor. The reverse bias voltage is applied through the p^+ implanted strips. The n^+ layer prevents the depletion region to reach the back plane.

The sensors in the SCT are on average 289.5 μm thick and are fully depleted at around 65 V [80]. To reduce noise and provide fast readout, the initial voltage at which the sensors operate is 150 V. After a few years of exposure to radiation, the n-type silicon will gradually invert into p-type silicon, leaving the sensors with a (p^+, p, n^+) structure. As a result, the p-n junction will move to the opposite side of the sensor and higher voltages, up to 350 V, will be required to achieve full depletion.

Throughout the lifetime of the SCT, it is very important to monitor the depletion depth since it affects the charge collection efficiency and spatial resolution of the sensors, and consequently, the quality of the recorded data.

5.2 Measurements using Cosmic Data

Two different techniques are explored to measure the depth of the depletion region of the SCT silicon sensors. For these measurements, cosmic-ray data taken between 2008 and 2009 are used and, given the angle of incidence of the tracks, only the barrel region is considered. During this period, the SCT operated in different configurations. Amongst them, the ones of particular interest to the depletion depth studies are the data samples where the SCT

modules were biased with voltages different from the nominal 150 V, since with them the effect of under-depletion can be characterised. Table 5.1 lists the cosmic-ray data samples used in this analysis.

Solenoid	Bias Voltage [V]
OFF	150 50
ON	300, 250, 200, 150, 100 60, 50, 40, 30

Table 5.1: Bias voltage and status of the solenoid magnetic field for the cosmic-ray data samples used in the depletion depth measurements.

5.2.1 Track Selection

Tracks are included in the analysis if they have:

- at least 8 hits in the silicon detectors *or* at least 30 hits in the TRT,
- a transverse impact parameter $|d_0| \leq 500$ mm,
- a transverse momentum $p_{\mathrm{T}} \geq 1$ GeV,
- a TRT event phase[1] between -10 and 40 ns.

5.2.2 Track Depth Approach

This first method is purely geometrical and uses information from each hit-on-track to determine the depletion depth. A similar approach was used in [81] for irradiated ATLAS Pixel detector sensors.

[1]The event phase is defined as the time within a readout window when a cosmic-ray track passes through the TRT. In collision events, the event phase is replaced with the LHC clock.

Description

In figure 5.3 a simple schematic representation of the plane transverse to the direction of the strips (xy-plane) of an SCT sensor is presented. If a charged particle enters the detector in x_0, with an incidence angle φ, then as it moves through, it liberates charge that must travel a distance d_i towards the silicon strips. Having a binary readout, the information of a strip will only be transferred if it collects enough charge above the 1 fC threshold. The group of fired strips is called a cluster and its size depends on the depth D of the depletion layer.

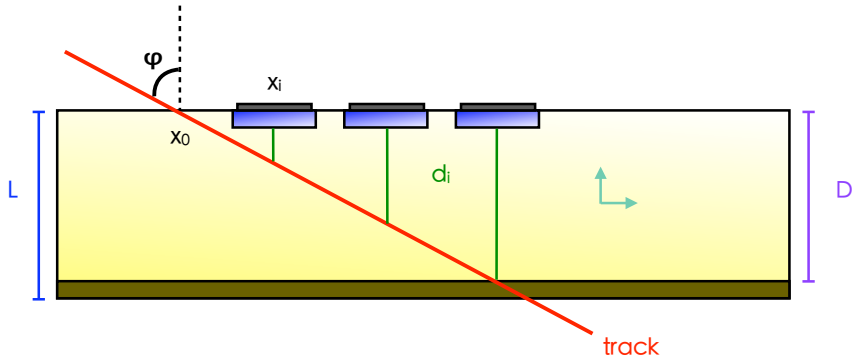


Figure 5.3: Schematic view of a particle crossing an SCT sensor from the plane transverse to the direction of the strips.

The track depth d_i, the distance between the track and the i-th strip, is given by

$$d_i = \frac{x_i - x_0}{\tan\varphi}, \tag{5.1}$$

where x_i is the local position of each fired strip. The incidence angle φ is measured from the normal to the sensor surface and, by definition, is positive in the clockwise direction. The entry point x_0 is measured at the surface of the sensor and, since the local reference frame is defined in the central plane

between the two active silicon layers [82], it is calculated as

$$x_0 = x_c + \frac{D}{2}\tan\varphi. \tag{5.2}$$

The first term in equation (5.2), x_c, corresponds to the extrapolated cluster position and is obtained by subtracting the track residual[2] from the measured cluster position, which basically gives the x position of the track in the middle plane.

Since D is the quantity we want to measure, as a first approximation it can be replaced with the sensor thickness L, so that

$$x_0' = x_c + \frac{L}{2}\tan\varphi, \tag{5.3}$$

which introduces a systematic displacement of the entry point

$$\Delta x_0 = \frac{L-D}{2}\tan\varphi. \tag{5.4}$$

The track depth is computed for every strip in every cluster and stored in a histogram. The falling edge of the resulting distribution is fitted with an error function given by

$$f(D_f - x) = 1 - \frac{a}{2}\mathbf{Erfc}\Big(\frac{D_f - x}{b\sqrt{2}}\Big) = 1 - \frac{a}{\sqrt{\pi}}\int\limits_{\frac{D_f-x}{b\sqrt{2}}}^{\infty}\exp(-t^2)dt, \tag{5.5}$$

where a, b and D_f are the parameters extracted from the fit. D_f denotes the inflection point of the function and is associated to the depletion depth.

[2]The track residual is the shortest distance between the fitted track and the hit's centre of gravity.

Selection

At the cluster level additional requirements are imposed. First of all, the absolute value of the local incidence angle φ should be less than 60°; at large incidence angles, the 1 fC threshold plays a more crucial role, affecting the cluster size by producing split clusters. To avoid the edges of the sensors, the calculated entry point x_0 is required to be within ±25 mm from the origin[3] (in total, the sensors are $\sim$ 60 mm wide). Finally, the extrapolated cluster position x_c must lie between the first and last strip of the cluster.

Preliminary Results

Figure 5.4 shows the track depth distribution for cosmic-ray data in the absence of the solenoid magnetic field and with the sensors biased at 150 V. Two interesting features stand out from this distribution: a slight dip close to 0 mm and a long tail reaching depths of 1 mm, even though the sensors are only 0.285 mm thick on average.

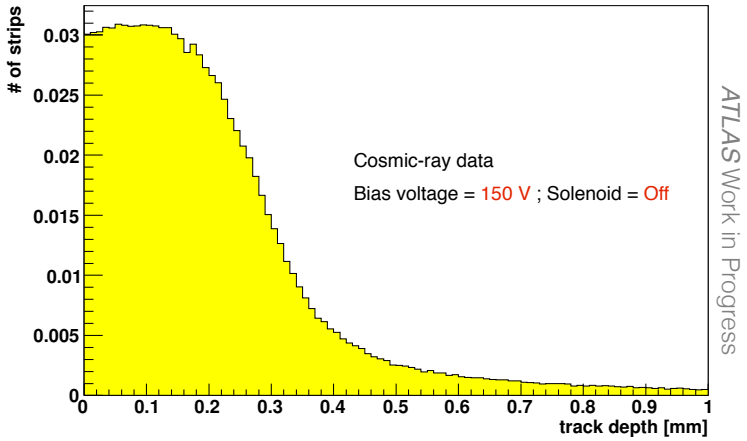


Figure 5.4: Track depth distribution for cosmic-ray data with 150 V bias voltage and with the solenoid magnetic field off.

[3]The origin of the local reference frame is located in the geometrical centre of the SCT sensors.

By dividing the sample in 10° angle slices, as illustrated in figure 5.5, it becomes clear that the dip is more pronounced at large incidence angles while the tail dominates the small angle regime. Looking at equation (5.1), the track depth is inversely proportional to the tangent of the local incidence angle, so if the angle is small, then the tangent will be small and consequently the track depth will be large. Large angles are affected by split clusters. Incidence angles between 20° and 30° produce a flat distribution around 0 mm and almost no tail extending to very large depths. All samples show the same behaviour so this interval will be used to determine the depletion depth in all cases.

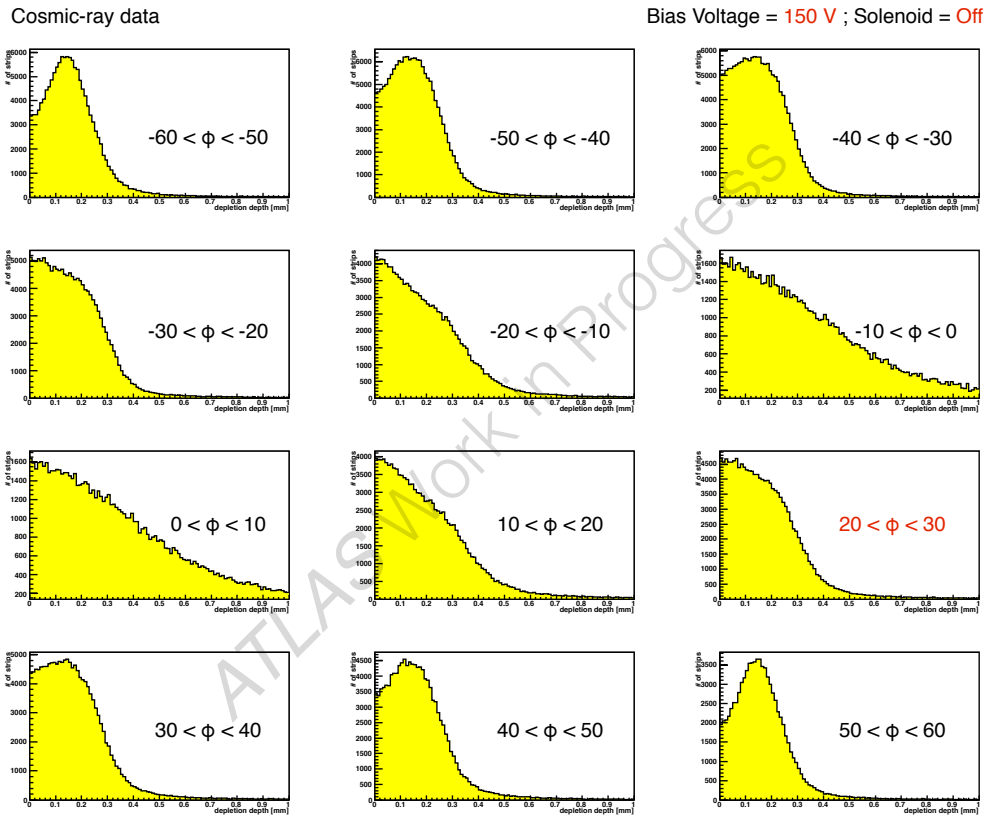


Figure 5.5: Track depth distribution in different incidence angle slices for cosmic-ray data at 150 V bias voltage and with the solenoid magnetic field off.

In figure 5.6, the fitted (using equation (9.1)) track depth distributions for

two cosmic-ray data samples, 150 V and 50 V of bias voltage, taken with the solenoid magnetic field off are shown along with the resulting fit parameters. Fitting ranges with similar $\chi^2/\mathrm{N_{dof}}$ showed a variation of 4-5% in the measured value of D_f. For the 150 V case, the value for D_f is consistent with the sensor thickness while for the 50 V case, it suggests that the sensors are slightly under-depleted.

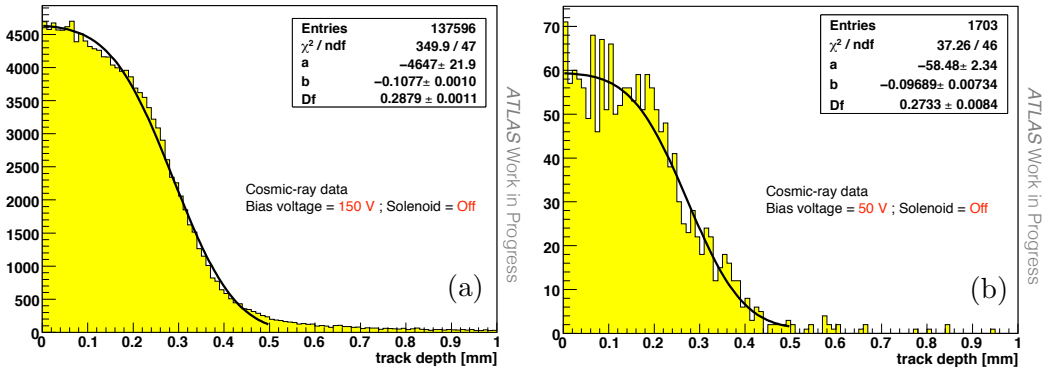


Figure 5.6: Fitted track depth distributions for cosmic-ray data with (a) 150 V and (b) 50 V bias voltage and the solenoid magnetic field off.

Figure 5.7 shows the fitted track depth distributions for the cosmic-ray data samples taken with the presence of the solenoid magnetic field and with the sensors biased at a voltage smaller than the depletion one.

When the bias voltage is smaller than the depletion voltage, the assumption made in equation (5.3) does not hold since $D \neq L$ and so the value of D_f has to be corrected for this effect. From equation (5.4) it can be inferred that the difference between the measured and the true value of the depletion depth is

$$D_f - D = \frac{L - D}{2}, \tag{5.6}$$

which means that the correction factor for the systematic displacement of the entry point is given by

$$D = 2D_f - L. \tag{5.7}$$

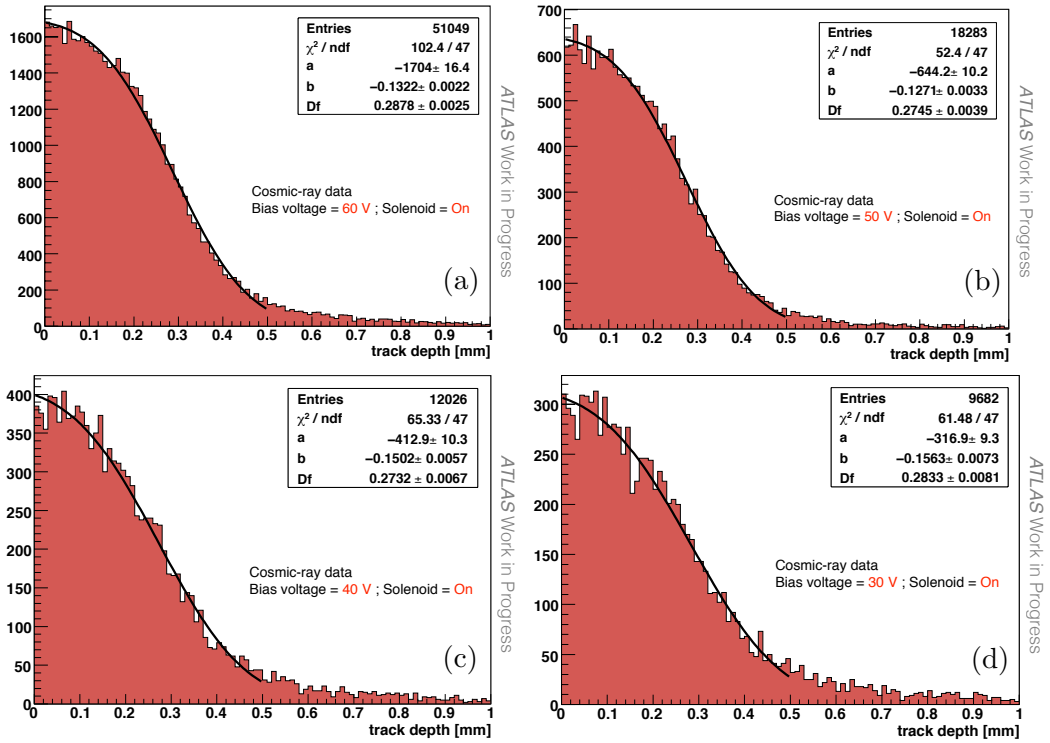


Figure 5.7: Fitted track depth distributions for cosmic-ray data with (a) 60 V, (b) 50 V, (c) 40 V and (d) 30 V bias voltage and the solenoid magnetic field on.

The values for D_f and D for the different samples are summarised in table 5.2. Figures 5.8 and 5.9 show the corrected value of the depletion depth as a function of the square root of the bias voltage for the samples without and with the solenoid magnetic field, respectively. In these plots, the black markers correspond to the data measurements and the solid red line to the expected values of the depletion depth if

$$D = L\sqrt{\frac{V_b}{V_d}} \tag{5.8}$$

with the sensor thickness $L = 289.5\mu$m, the depletion voltage $V_d = 64.8$ V and with V_b the bias voltage.

Solenoid	Bias Voltage [V]	D_f [mm]	D [mm]
OFF	150	0.2879 ± 0.0011	0.2879 ± 0.0011
	50	0.2733 ± 0.0084	0.2571 ± 0.0168
ON	60	0.2878 ± 0.0025	0.2861 ± 0.0050
	50	0.2745 ± 0.0039	0.2595 ± 0.0078
	40	0.2732 ± 0.0067	0.2569 ± 0.0134
	30	0.2833 ± 0.0081	0.2771 ± 0.0162

Table 5.2: Calculated values for the depletion depth for the different bias-voltage cosmic-ray data samples.

For data taken in the absence of the magnetic field, this method shows reasonable agreement with the expected values. This is not the case for the data immersed in the magnetic field. Being a geometrical method, this approach is not able to reproduce the effects of the magnetic field on the electron-hole pairs drifting in the silicon due to the Lorentz Force.

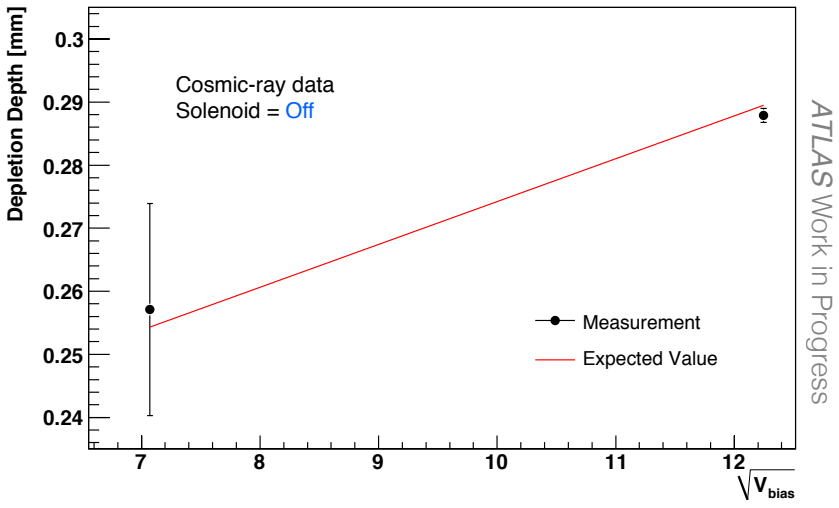


Figure 5.8: Comparison between the measured depletion depth and the expected values for samples without the solenoid magnetic field.

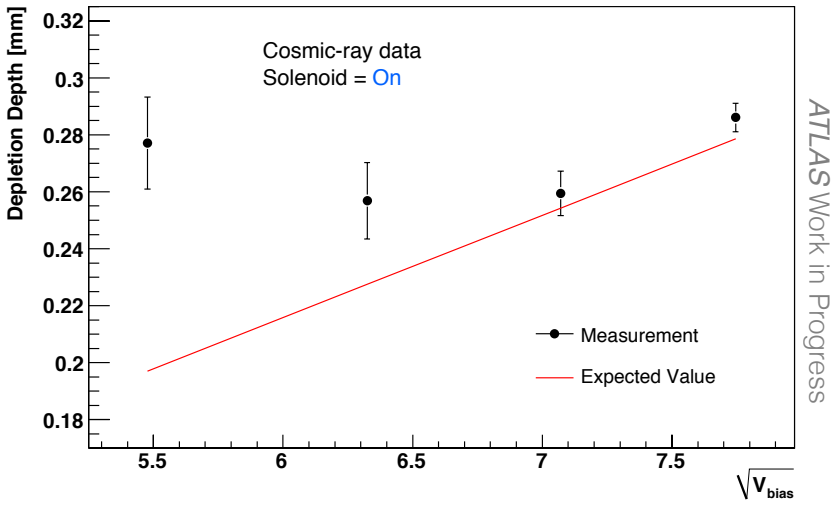


Figure 5.9: Comparison between the measured depletion depth and the expected values for samples with the solenoid magnetic field.

5.2.3 Fit to Slope Approach

Description

The distribution of the average cluster size as a function of the local incidence angle φ, is sensitive to changes in the bias voltage at which the sensors are set. When a sensor is not fully depleted, some of the electron-hole pairs liberated by a traversing charged particle recombine with the free carriers, which means that, for a given incidence angle, on average there will be less strips that collect enough charge above threshold. By monitoring the slope of the average cluster size distribution with respect to φ, the depletion depth of the SCT sensors can be indirectly estimated. A variety of functions can be fitted to obtain the value of the slope but the simplest one, a straight line, is chosen for this measurement between the range $0° < \varphi < 20°$; including the minimum of the distributions in the fit, which is associated to the Lorentz angle, results in large fluctuations depending on the chosen fitting range.

Preliminary Results

Figure 5.10 shows the average cluster size vs local incidence angle distribution for cosmic-ray data samples, collected during 2009, at different bias voltages and with the solenoid magnetic field on. For bias voltages greater than the depletion voltage, all distributions lie on top of each other, while for bias voltages lower than the depletion voltage the distributions get distorted and the average cluster size decreases.

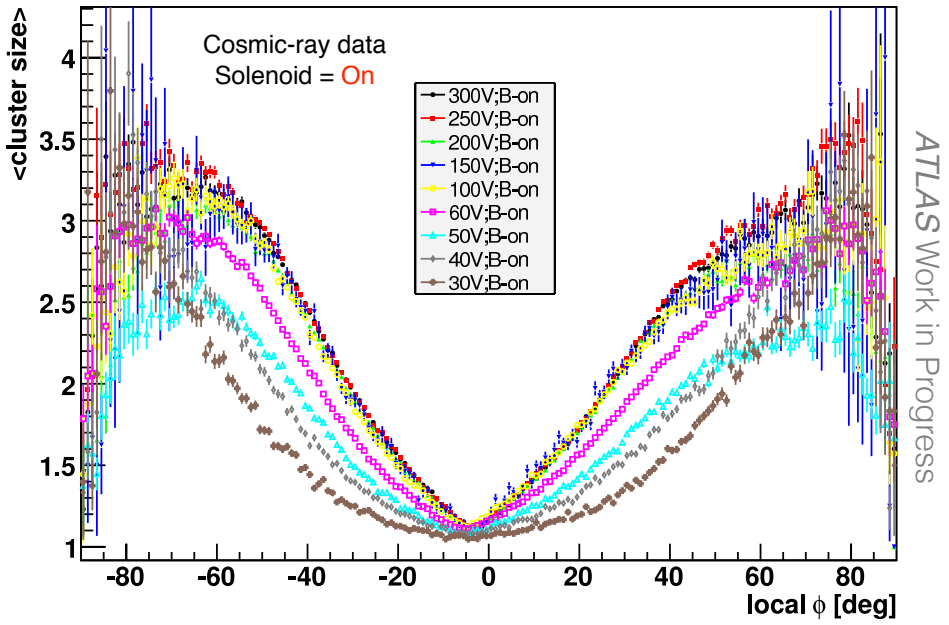


Figure 5.10: Average cluster size as a function of the local incidence angle φ for different bias voltages with the solenoid magnetic field on.

In figure 5.11, examples of the linear fits done to the cluster size vs φ distributions for different bias voltages are presented. The slope values extracted from the linear fits, listed in table 5.3, are plotted as a function of the square root of the bias voltage in figure 5.12. If a straight line is fitted to the plateau region and another one to the low voltage region, it can be seen that the two intersect at around $\sqrt{V_{bias}} = 8.5$, approximately the value of the depletion voltage.

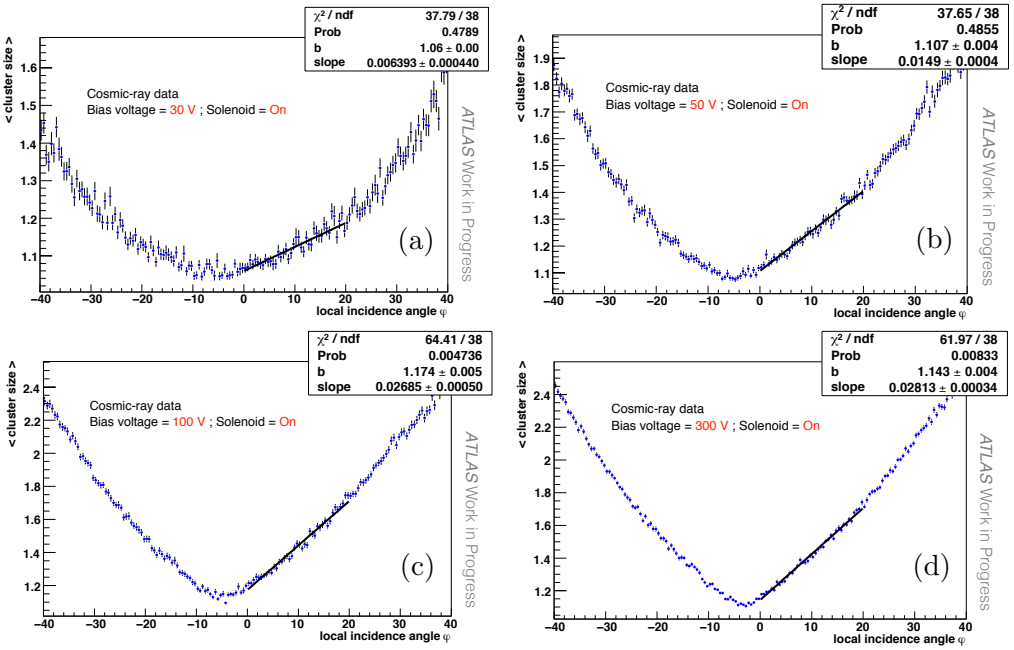


Figure 5.11: Examples of linear fits to the average cluster size vs incidence angle distributions for cosmic-ray data taken with the solenoid magnetic field on and with bias voltages of (a) 30 V, (b) 50 V, (c) 100 V and (d) 300 V.

Bias Voltage [V]	slope
300	0.0281 ± 0.0003
250	0.0278 ± 0.0003
200	0.0274 ± 0.0004
150	0.0261 ± 0.0006
100	0.0268 ± 0.0005
60	0.0203 ± 0.0003
50	0.0149 ± 0.0004
40	0.0105 ± 0.0005
30	0.0064 ± 0.0004

Table 5.3: Fitted slope values for the different bias-voltage cosmic-ray data samples.

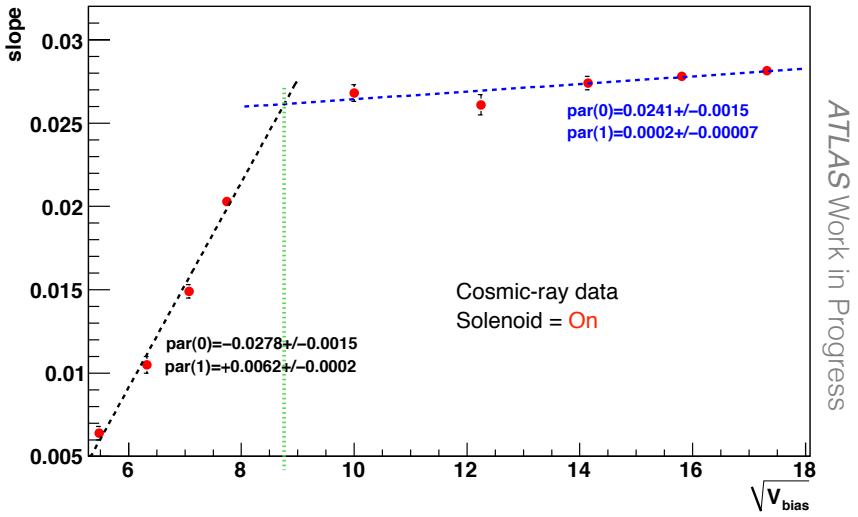


Figure 5.12: Fitted value of the slope as a function of the square root of the bias voltage. The intersection of the fitted lines is close to the depletion voltage.

5.2.4 Conclusion

This preliminary study of the depletion depth of the SCT sensors using cosmic-ray data shows that two approaches can be used to monitor this quantity, one valid only in the absence of the solenoid magnetic field and the second one which measures it indirectly. The "Fit to Slope" approach can be used with or without the magnetic field and has the additional advantage of being relatively simple to implement in the SCT data quality monitoring scripts. If the value of the slope of the average cluster size vs φ distribution goes below a certain value (below the plateau region), it will be an indication that the sensors are not fully depleted and that the bias voltage needs to be increased for an optimal performance.

Part III

Analysis

Chapter 6

Event Selection & Reconstruction Efficiencies

6.1 Minimum Bias in ATLAS

As discussed in section 1.3, the definition of minimum bias depends on the trigger requirements chosen by each experiment. A common way to measure the properties of minimum bias events is to use data collected with a double-sided trigger to suppress the contributions from single-diffraction (also known as Non-Single Diffractive), as was done for example by [83]. This approach introduces model dependent corrections for the diffractive components. ATLAS has adopted a slightly different philosophy [84]: to perform an inelastic measurement that is easily reproducible thanks to a well defined final state selection and which contains minimal model dependence.

In ATLAS, two complementary trigger strategies for selecting minimum bias events are available [85, 86]. The measurement presented in this thesis is done using data selected with the MBTS and BPTX triggers which are described below.

6.1.1 MBTS

The signal for this trigger is read from Minimum Bias Trigger Scintillators which are located at ±3.56 m from the centre of ATLAS. Each MBTS wheel is segmented into 16 plates of trapezoidal shape, called counters, arranged in two rings in pseudorapidity, covering the range $2.1 < |\eta| < 3.8$, and 8 sections in azimuth (figure 6.1). The MBTS wheels are mounted on the cryostat of the LAr calorimeter, perpendicular to the beam direction, and the scintillators are connected to PhotoMultiplier Tubes (PMTs) of the Tile calorimeter which results in a fast readout.

The sensitivity of the MBTS to minimal detector activity proves useful for the early commissioning stages. After a few months of operation at high luminosity, it will suffer from radiation damage, and a different strategy will be needed for minimum bias measurements.

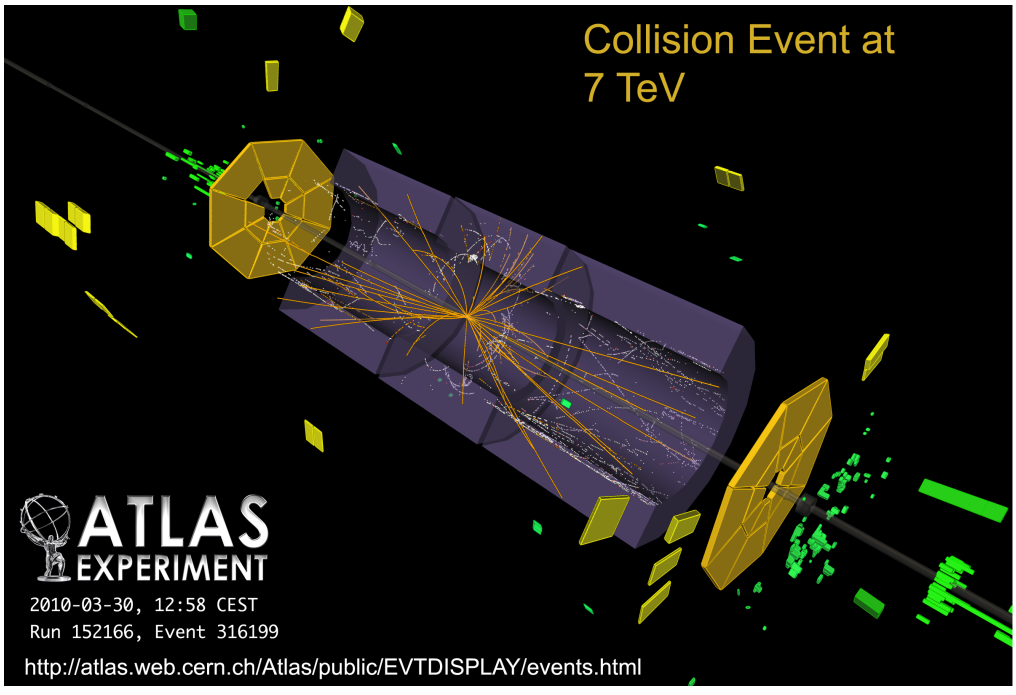


Figure 6.1: Display of a 7 TeV collision event, recorded in March 2010 by the ATLAS detector, which triggered the MBTS counters, highlighted in orange. From [49].

An event is selected if the energy deposited in the counters is above a discriminator threshold of 50 mV [87]. The counter multiplicity is measured independently for each side. These requirements are grouped into different trigger setups, or menus, that are loaded at Level 1. Table 6.1 summarises the thresholds used for three of the most important trigger menus for early data-taking.

Trigger Menu	Requirements
L1_MBTS_1	≥ 1 counter above threshold on **either** side of the detector
L1_MBTS_1_1	≥ 1 counter above threshold on **both** sides of the detector
L1_MBTS_2	≥ 2 counters above threshold **in total**

Table 6.1: MBTS trigger menus at Level 1.

6.1.2 BPTX

The ATLAS BPTX stations [88] are located at both sides of the detector, 175 m away from the interaction point. Each station has four electrostatic beam pick-up detectors which are arranged symmetrically in the direction transverse to the beam pipe.

The BPTX have two purposes:

- Monitor the position and structure of the beams and provide timing signals to allow ATLAS to synchronise its front-end electronics with the LHC collision clock.

- Provide a trigger condition to indicate that bunches are passing through ATLAS. The signal from the BPTX is discriminated and shaped into 25 ns long pulses that are given to the Central Trigger Processor (CTP) to provide *filled bunch triggers*. The BPTX triggers are usually combined with others as a confirmation of beam crossings.

Figure 6.2 shows the BPTX pulses produced by a proton bunch that circulated 8 times around the LHC on September 2008. The pulses are separated by 89 μs and after each turn, the intensity of the signal degrades due to losses and de-bunching of the beam.

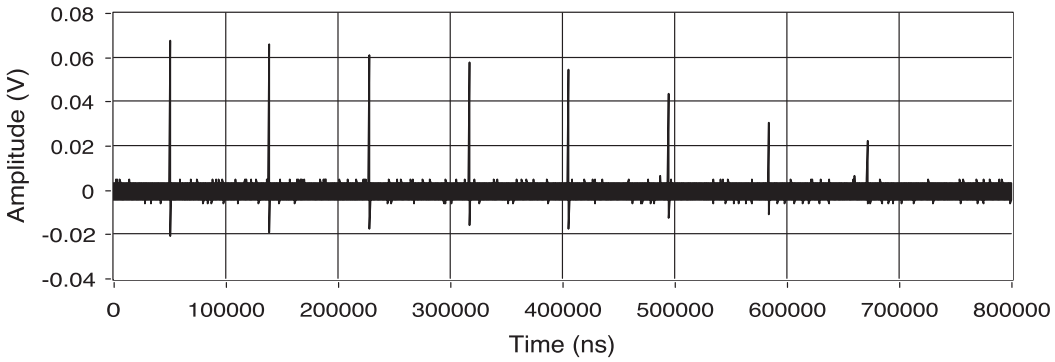


Figure 6.2: BPTX signal produced by a proton bunch passing through ATLAS eight times. From [88].

6.2 Event and Track Selection Requirements

The analysis of two-particle angular correlations is done on Minimum Bias charged-track data collected during 2009 and 2010, at two centre-of-mass energies. It uses the full 900 GeV data sample, which corresponds approximately to an integrated luminosity of 7 μb^{-1}, and the first 190 μb^{-1} of 7 TeV data. Early data consists of low instantaneous luminosity runs where the probability of additional interactions in the same bunch crossing is of the order of 0.1%. The event and track selection criteria is identical to the one used in the ATLAS charged-particle multiplicity analysis, more commonly known as Minimum Bias analysis [84].

Only events that occurred when the ATLAS Inner Detector was fully operational and the solenoid magnetic field was at its nominal strength of 2 Tesla,

are used. To reduce the contribution from background events (e.g. beam-gas interactions or beam-halo) and from non-primary tracks, events are also required

- to have been triggered by the single-arm, single-counter Level 1 MBTS (L1_MBTS_1 trigger menu) and by a coincidence of BPTX signals between the two sides of ATLAS,
- to contain at least one primary vertex (the one with the highest $\sum p_{\mathrm{T}}^2$) reconstructed using the *beam spot*[1] (BS) information and at least two tracks with:
  - $p_{\mathrm{T}} > 100$ MeV,
  - $|d_0^{BS}| < 4$ mm,
  - transverse and longitudinal errors of $\sigma(d_0^{BS}) < 5$ mm and $\sigma(z_0^{BS}) < 10$ mm, respectively,
  - at least 1 hit in the Pixel detector and 4 hits in the SCT,
  - at least 6 hits in the silicon detectors (Pixel + SCT)
- not to have a second interaction vertex associated to more than four tracks in order to remove *pile-up*, events with more that one proton-proton interaction in a single bunch-crossing,
- to contain a minimum number of tracks n_{sel}, reconstructed with either the Inside-Out or Low-pt tracking algorithms, with:
  - $p_{\mathrm{T}} > 100$ MeV,

[1]The beam spot is defined as the luminous region produced by the pp collisions. For the Minimum Bias analysis, the beam spot parameters (position, size and tilt) are calculated using a maximum-likelihood fit to the distribution of reconstructed primary vertices [89].

- $|\eta| < 2.5$ (inside the acceptance of the ID),
- at least 1 hit in the b-layer of the Pixel detector (only if the extrapolated track crossed an active region of a Pixel module that was not disabled),
- at least 1 hit in any of the layers of the Pixel detector,
- at least 2 ($p_{\mathrm{T}} > 100$ MeV), 4 ($p_{\mathrm{T}} > 200$ MeV) or 6 ($p_{\mathrm{T}} > 300$ MeV) hits in the SCT,
- $|d_0^{PV}| < 1.5$ mm (PV - measured with respect to the primary vertex),
- $|z_0 \sin\theta^{PV}| < 1.5$ mm,
- a track-fit χ^2 probability of at least 0.01 if the track has a $p_{\mathrm{T}} > 10$ GeV (to remove tracks with mis-measured high-p_{T} due to misalignment or interactions with the detector material)

The impact parameter cuts and the requirement of at least one hit in the Pixel detector help discriminate between non-primary and primary tracks. The different number of SCT hits as a function of p_{T} helps to reduce the fraction of badly reconstructed tracks; a short track length translates into worse momentum resolution and migration effects in which low-p_{T} tracks are reconstructed as having very high momentum. Studies showed that a fraction of tracks with mis-measured p_{T} survived after applying the silicon hit requirements [84]. Most of these tracks ($p_{\mathrm{T}} > 20$ GeV), however, do not have a good quality fit so they can be rejected by applying a cut on the χ^2 probability.

At 7 TeV, two different phase-space regions with varying contributions from diffractive events are explored: $n_{ch} \geq 2$ and $n_{ch} \geq 20$, where n_{ch} is the charged-particle multiplicity. For the measurement at 900 GeV, only the multiplicity cut of $n_{ch} \geq 2$ is used due to limited statistics at high multiplicity.

The total number of events and tracks satisfying the above requirements for each phase-space region are shown in table 6.2.

Track Multiplicity Cut	$\sqrt{s}$ [TeV]	Selected Events	Selected Tracks
$n_{sel} \geq 2$	0.9	357,523	4,532,663
	7	10,066,072	209,809,430
$n_{sel} \geq 20$	7	4,029,565	153,553,766

Table 6.2: Total number of selected events and tracks in 900 GeV and 7 TeV data. All the different phase-space regions are defined for a $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$.

6.3 Reconstruction Efficiencies

The event selection and reconstruction is affected by the resolution of the detector and the performance of the reconstruction algorithms. In order to correct for these effects, knowledge of the reconstruction efficiencies is necessary. This section describes how the efficiencies associated to tracking, vertexing and triggering relevant for this analysis were determined. These distributions were derived by the Minimum Bias analysis in [84, 90].

6.3.1 Tracking

The track reconstruction efficiency, ϵ_{trk}, is determined from Monte Carlo. The same event and track selection criteria as in data are applied (section 6.2). The first step in the determination of the efficiency consisted in making sure that the distributions of track properties reconstructed in data were correctly reproduced by the simulation [84, 91].

The efficiency, parameterised in bins of p_{T} and η, is defined as

$$\epsilon_{trk}(p_{\mathrm{T}}, \eta) = \frac{N_{rec}^{matched}(p_{\mathrm{T}}, \eta)}{N_{gen}(p_{\mathrm{T}}, \eta)} \tag{6.1}$$

where $N_{rec}^{matched}(p_{\mathrm{T}}, \eta)$ corresponds to the total number of reconstructed tracks that are matched to a generated charged particle and $N_{gen}(p_{\mathrm{T}}, \eta)$ is the total number of generated charged particles in the given (p_{T}, η) bin.

The matching of reconstructed tracks to generated particles is carried out using a cone+hit strategy. Only primary particles, defined at the generator level as either having a lifetime greater than 3×10^{-11} seconds or being the decay products of a particle with a lifetime less than 3×10^{-11} seconds, enter the matching algorithm.

The cone-matching, done in the $\eta\phi$-plane, associates a charged particle with the track that has the smallest $\Delta\mathcal{R} = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ with respect to it, within a cone radius of 0.15. In some cases, the track will be matched to a charged particle that happened to be nearby but did not produce such track. The contribution from these fake tracks can be reduced by requiring that the trajectory of the charged particle and the track share a hit in the Pixel detector. The remaining un-matched tracks are considered secondaries.

Figures 6.3 and 6.4 show the p_{T} and η projections of the track reconstruction efficiencies at 900 GeV and 7 TeV, respectively. The error bands correspond to the total uncertainty in each bin (a discussion of the uncertainties on this measurement is deferred to section 8.2.2). The small differences between the two energies arise from the different configuration and number of disabled modules in the silicon detectors between the two periods of data-taking. In both cases, however, the tracking efficiency drops for values of $|\eta| > 1.0$, because of the larger amount of material that particles have to go through in this region (end-cap), and there is a turn-on as a function of p_{T} related to the constraints on the number of silicon hits.

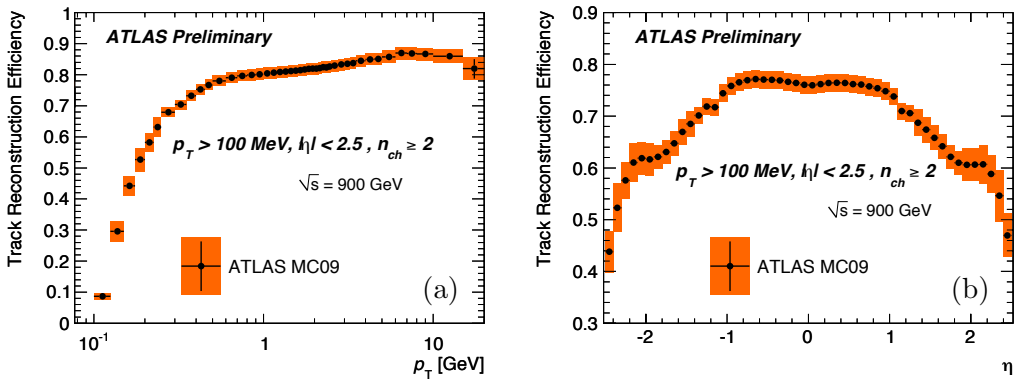


Figure 6.3: Track reconstruction efficiency at 900 GeV as a function of (a) p_{T} and (b) η for tracks with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$. From [90].

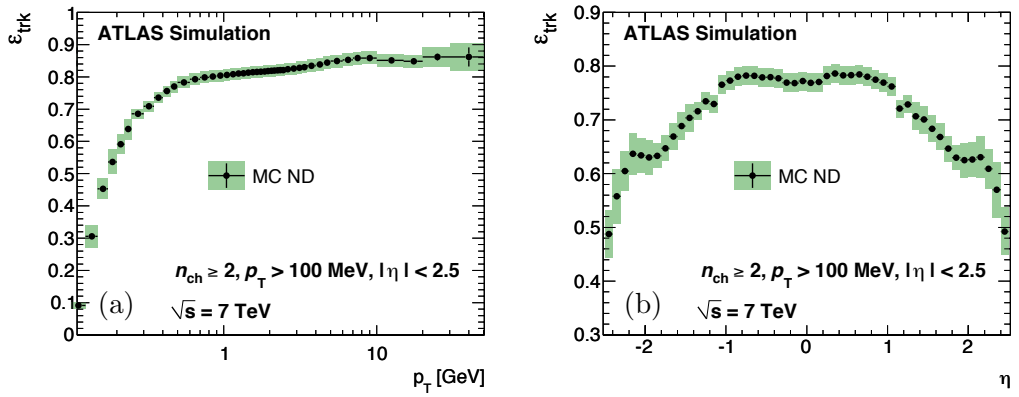


Figure 6.4: Track reconstruction efficiency at 7 TeV as a function of (a) p_{T} and (b) η for tracks with $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$. From [84].

6.3.2 Vertexing

The vertex reconstruction efficiency, ϵ_{vtx}, is derived from data. The event and track selection described in section 6.2 is applied with the exception of the requirements on the presence of a primary vertex, and hence, the transverse and longitudinal impact parameter cuts with respect to it. Instead, tracks are selected if they have a $|d_0^{BS}| < 1.8$ mm. The efficiency, as a function of the beam spot track multiplicity n_{sel}^{BS}, is given by:

$$\epsilon_{vtx}(n_{sel}^{BS}) = \frac{N_{vtx}(n_{sel}^{BS})}{N(n_{sel}^{BS})} \tag{6.2}$$

where $N(n_{sel}^{BS})$ is the total number of triggered events satisfying these modified selection criteria and $N_{vtx}(n_{sel}^{BS})$ the fraction of these events containing a reconstructed vertex. The efficiencies for both energies are shown in figure 6.5. The error bands correspond to the total uncertainty in each bin (see section 8.2.2). At both energies, the efficiency rapidly reaches 100% after a 90-92% value in the first multiplicity bin. For events with $n_{sel}^{BS} = 2$, the efficiency shows a dependence on the separation in z_0^{BS} between the tracks; at low n_{sel}^{BS}, it is more likely to reconstruct a primary vertex if the tracks are close together. In figures 6.6 and 6.7, the vertex efficiency distributions as a function of Δz_0^{BS} for events containing at least one track with a $100 < p_{\mathrm{T}}^{min} < 200$ MeV and $p_{\mathrm{T}}^{min} > 200$ MeV are shown for 900 GeV and 7 TeV, respectively.

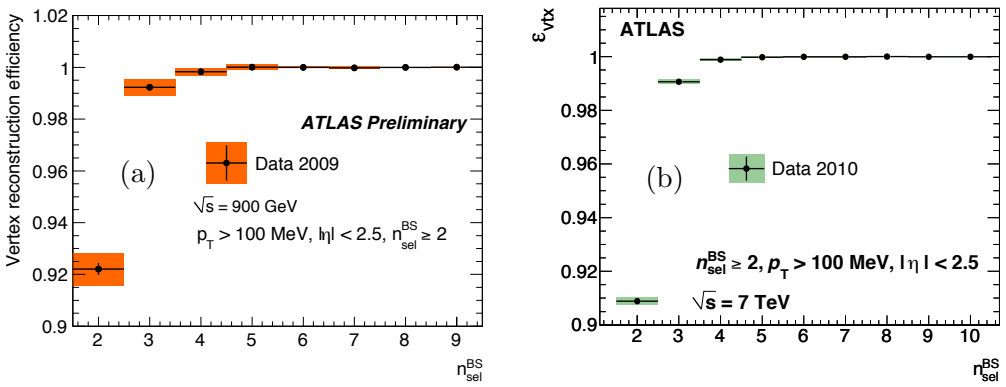


Figure 6.5: Vertex reconstruction efficiency, as a function of n_{sel}^{BS}, at (a) 900 GeV [90] and (b) 7 TeV [84].

6.3.3 Triggering

The trigger reconstruction efficiency, ϵ_{trig}, is also calculated from data. A control trigger is used to select events from random filled bunch crossings, filtered at the HLT. For the 900 GeV case, the L2 requires at least 7 clusters in the Pixel detector plus a minimum of 7 SCT hits and 1 track with $p_{\mathrm{T}} > 200$ MeV at the EF level. For 7 TeV, only 4 Pixel clusters and 4 SCT hits are

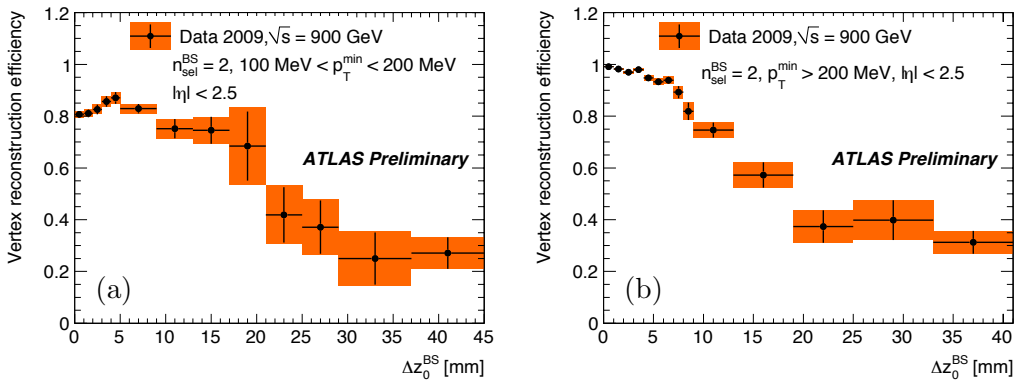


Figure 6.6: Δz_0^{BS} dependence of ϵ_{vtx} at 900 GeV for events with at least one track with (a) $100 < p_{\mathrm{T}}^{min} < 200$ MeV and (b) $p_{\mathrm{T}}^{min} > 200$ MeV. From [90].

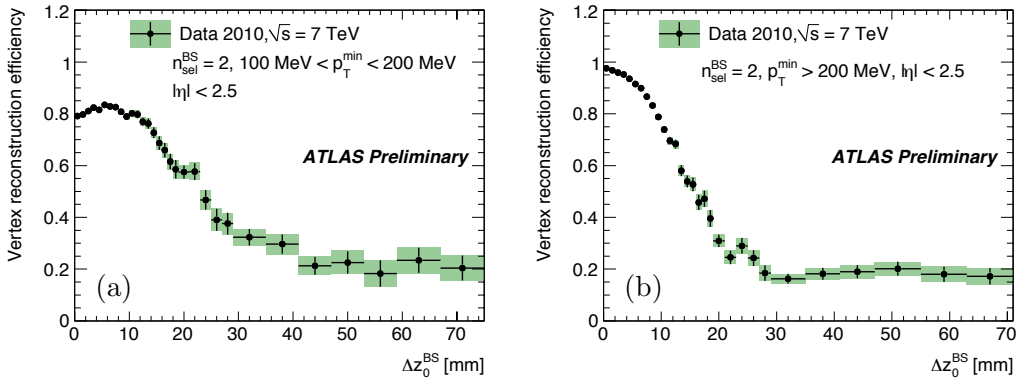


Figure 6.7: Δz_0^{BS} dependence of ϵ_{vtx} at 7 TeV for events with at least one track with (a) $100 < p_{\mathrm{T}}^{min} < 200$ MeV and (b) $p_{\mathrm{T}}^{min} > 200$ MeV. From [90].

required, with no additional input from the EF.

The modified selection criteria described in section 6.3.2 are applied; the vertex requirement is dropped in this case due to possible correlations between the trigger and vertex efficiencies.

The trigger efficiency, as a function of n_{sel}^{BS}, is calculated as:

$$\epsilon_{trig}(n_{sel}^{BS}) = \frac{N_{rand+MBTS}(n_{sel}^{BS})}{N_{rand}(n_{sel}^{BS})} \tag{6.3}$$

with $N_{rand}(n_{sel}^{BS})$ the total number of selected events in the control sample and $N_{rand+MBTS}(n_{sel}^{BS})$ the fraction of these events that were also triggered by the

L1_MBTS_1. The efficiencies for both energies are presented in figure 6.8. The error bands correspond to the total uncertainty in each bin (see section 8.2.2). A 97% efficiency is observed for events with $n_{sel}^{BS} = 2$ and, as the multiplicity increases, it reaches 100%.

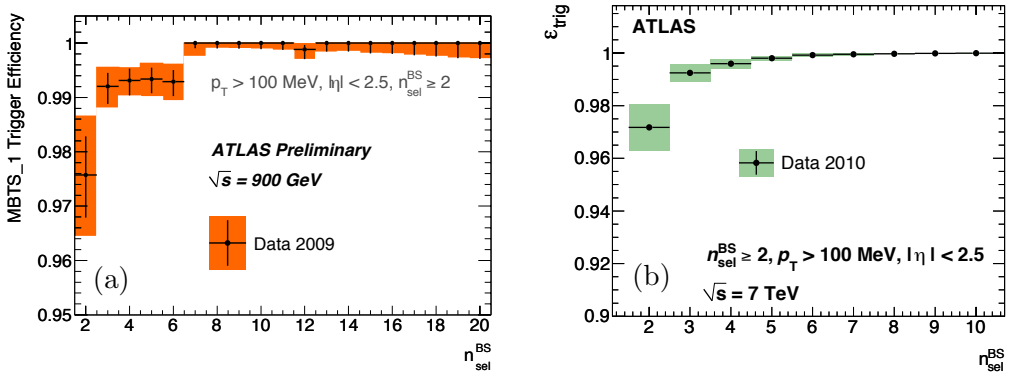


Figure 6.8: Trigger reconstruction efficiency, as a function of n_{sel}^{BS}, at (a) 900 GeV [90] and (b) 7 TeV [84].

Chapter 7

Correction Procedure & Closure Tests

7.1 Method

In order to disentangle the effects caused by the apparatus and the reconstruction algorithms from the true physics processes, the corrections described in the following sections are applied to the reconstructed data. The same procedure is followed for all distributions, both one- and two-dimensional, in all the regions of phase-space explored in this analysis.

7.1.1 Trigger and Vertex Selection

To account for the loss of events due to inefficiencies in the trigger and vertex requirements (section 6.2), an event-by-event weight, w_{ev}, given by

$$w_{ev} = \frac{1}{\epsilon_{vtx}(n_{sel}^{BS})} \frac{1}{\epsilon_{trig}(n_{sel}^{BS})} \tag{7.1}$$

is applied to both the foreground and multiplicity distributions, with $\epsilon_{vtx}(n_{sel}^{BS})$ and $\epsilon_{trig}(n_{sel}^{BS})$ the vertex and trigger reconstruction efficiencies, respectively,

as described in sections 6.3.2 and 6.3.3.

7.1.2 Probabilistic Track Removal

Due to inefficiencies in the track reconstruction, some tracks in the event can be lost. The effect that these missing tracks have on the correlations measurement is corrected for using the *Hit Backspace Once More* (HBOM) method [92]. The correction in this method is derived by repeatedly applying the (parameterisation of the) detector effects to the data and then extrapolating back to a detector effect of zero. This is a model-independent method that only requires as input the single track reconstruction efficiency as a function of the track p_{T} and η (section 6.3.1). The diagram in figure 7.1 illustrates the different steps of this correction procedure.

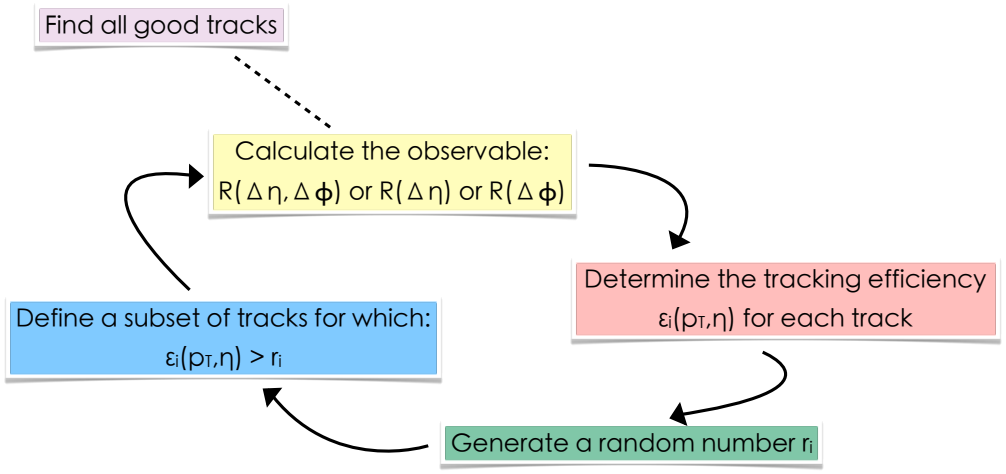


Figure 7.1: Probabilistic track removal method to correct for track reconstruction inefficiencies.

As a first step, the track-based observable, the correlation function, is computed using all the reconstructed tracks that satisfy the selection requirements described in section 6.2. Then, for each track, the track reconstruction efficiency $\epsilon_i(p_{\mathrm{T}}, \eta)$ is compared to a (unique) random number r_i, generated

uniformly between 0 and 1. If the random number is greater than the track reconstruction efficiency then the track is thrown away and removed from the sample.

Using a subset of the original tracks, defined by all those tracks for which

$$\epsilon_i(p_{\mathrm{T}}, \eta) > r_i, \tag{7.2}$$

the observable is computed again. This constitutes one iteration of the track removal procedure; the observable calculated using the full set of uncorrected tracks is defined as the *0th iteration*.

A second iteration takes the subset of tracks produced in the first iteration and uses a new set of random numbers to remove some of them according to their track reconstruction efficiency, and with the remaining tracks re-calculates the observable. Additional iterations are carried out in the same way. The choice of the number of track removal iterations is limited by statistics; there must be a sufficient number of events with sufficient tracks that some remain after all iterations. For the sample size used in this measurement, six iterations provide enough points to define a trend, and are not so many that the last iteration still contains enough tracks.

Figure 7.2 illustrates the effect of the track removal iterations on the pseudorapidity correlation function at 7 TeV for the track multiplicity cut $n_{sel} \geq 2$. The solid black circles correspond to $R(\Delta\eta)$ in the 0th iteration, using all selected tracks, and the different markers correspond to that same distribution after each iteration. It can be seen that by repeatedly applying the detector effects to the data, the correlation function is flattened out. This is related to the dependence of the correlation function on the track multiplicity, not only as a normalisation factor but also as a weight for each foreground pair (section 2.3).

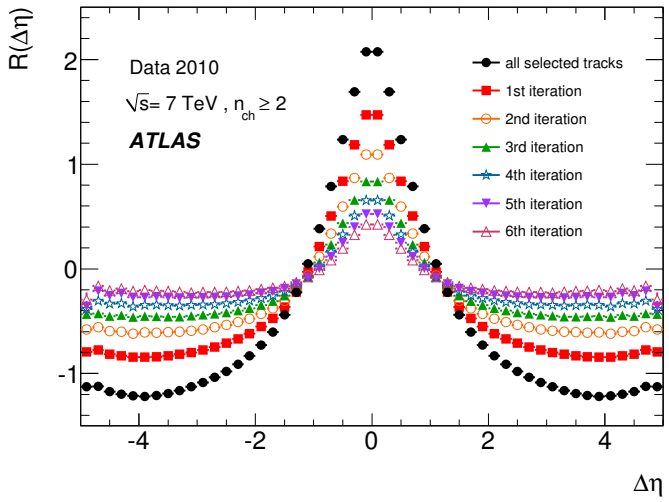


Figure 7.2: Two-particle pseudorapidity correlation function, $R(\Delta\eta)$, after each iteration of the probabilistic track removal correction method.

Finally, in order to obtain an estimate of the value of the correlation function at the particle level, without any detector effects, for each bin of the observable the values are plotted as a function of the iteration number (0, 1, 2, ..., 6) and the fit to the resulting distribution is extrapolated to -1. For this analysis, the fit and subsequent extrapolation use a third-degree polynomial as it is the simplest type of function that can fit all the bins of the observable. Figure 7.3 illustrates the polynomial fit for an inner and an outer $\Delta\eta$ bin of $R(\Delta\eta)$ at 7 TeV in the $n_{ch} \geq 2$ phase-space. The fits for all the bins in this distribution are shown in Appendix A.

Studies of the possible uncertainties introduced by the choice of the fitting function and the number of iterations used as input are described in section 8.2.3.

This probabilistic track removal correction method allows us to quantify the effect that detector inefficiencies have on an observable; in each iteration the detector effects are being re-applied to the data, which means that every time tracks are lost in exactly the same way as they are lost by the ATLAS

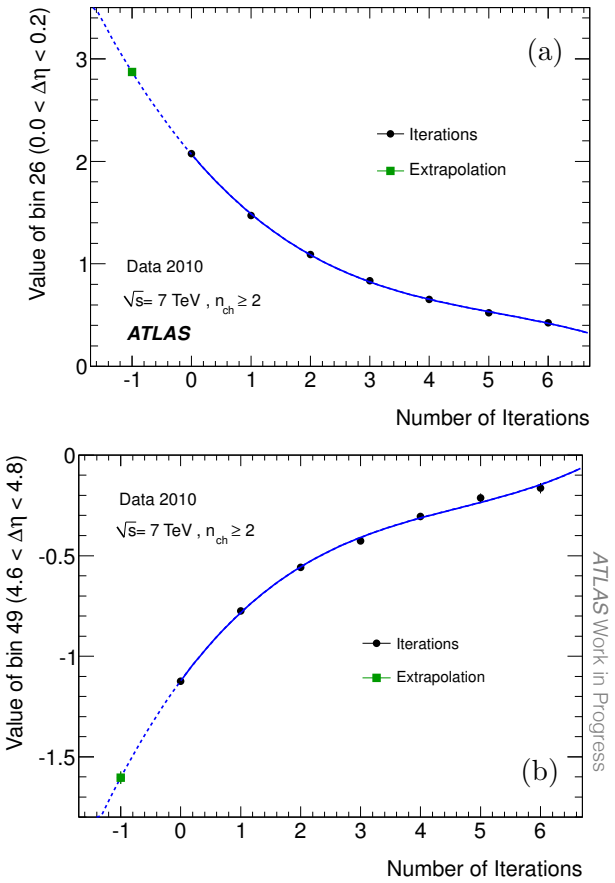


Figure 7.3: Example of third-degree polynomial fits to the values of $R(\Delta\eta)$ in (a) a central and (b) an outer bin.

detector. In other words, the 0th iteration corresponds to the effect of ATLAS, the first iteration to the effect of ATLAS squared, and so on. By extrapolating to the -1 iteration, the effect of ATLAS is removed and the true observable can be measured.

7.2 Monte Carlo Studies

7.2.1 Closure Tests

To test the effectiveness of the correction method, studies were done in Monte Carlo to check whether the generated, or *truth*, distributions were recovered

from the fully detector-simulated and reconstructed ones after applying the methodology described in the previous section (a procedure usually referred to as *closure tests*).

In order to eventually compare to data, an additional correction had to be applied to the Monte Carlo distributions to account for the fact that the beam spot position and width in the z-direction differed significantly between data and simulation. Not including this factor would imply a different vertex acceptance and slightly different kinematic distributions. For this reason, the simulation was re-weighted to reproduce the position of the primary vertex in the z-axis as seen in data (figure 7.4). Every MC distribution includes this event weight.

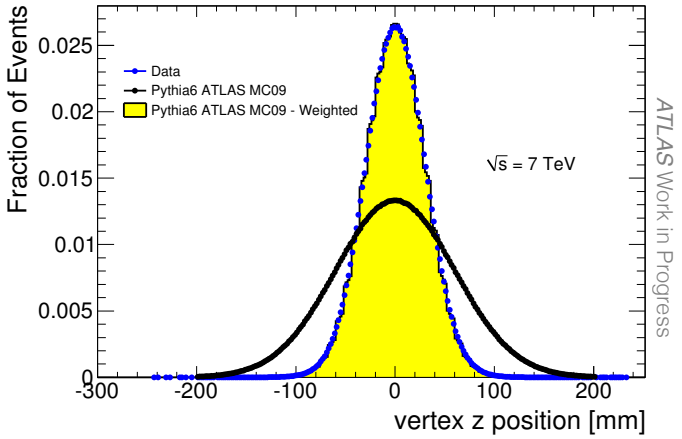


Figure 7.4: Distribution of the position of the primary vertex in the z-direction for data (blue markers) and the reconstructed PYTHIA6 MC09 Monte Carlo tune, before (black markers) and after (yellow histogram) re-weighting.

The pseudorapidity correlation function, $R(\Delta\eta)$, obtained by integrating the two-dimensional distribution over the range $0 < \Delta\phi < \pi$, will be used as an example to illustrate the closure of the correction method. For these tests, the ATLAS MC09 PYTHIA6 tune (described in section 1.4.3) is used. Different models and tunes were also studied with similar results (see discussion in section 8.2.1). Table 7.1 lists the total number of selected events and tracks

at both energies and regions of phase-space used for these Monte Carlo studies.

Track Multiplicity Cut	$\sqrt{s}$ [TeV]	Selected Events	Selected Tracks
$n_{sel} \geq 2$	0.9	790,877	8,746,908
	7	21,551,438	379,934,146
$n_{sel} \geq 20$	7	7,418,245	254,264,501

Table 7.1: Total number of selected events and tracks in 900 GeV and 7 TeV Monte Carlo MC09 PYTHIA6 tune. All the different phase-space regions are defined for a $p_{\mathrm{T}} > 100$ MeV and $|\eta| < 2.5$.

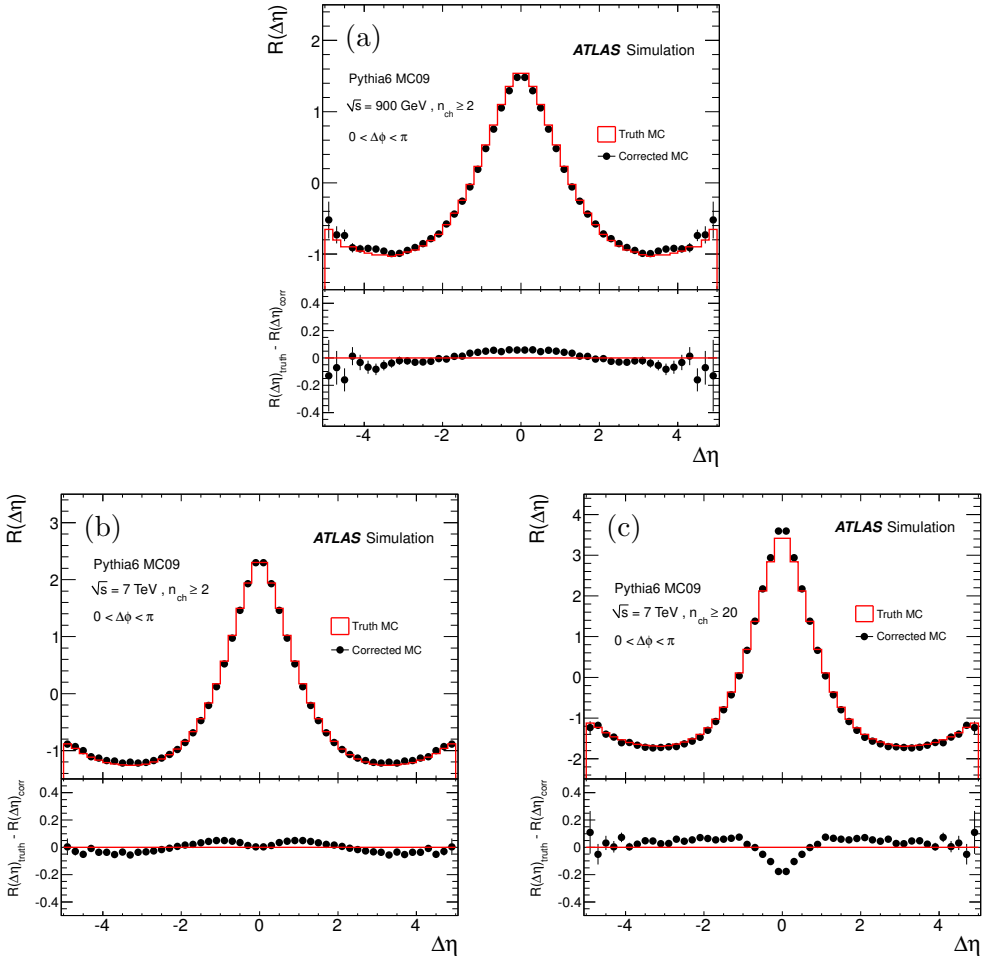


Figure 7.5: Comparison between the generated and corrected Monte Carlo pseudorapidity distributions, $R(\Delta\eta)$, for the PYTHIA6 MC09 tune at (a) 900 GeV and 7 TeV, for the phase-space regions (b) $n_{ch} \geq 2$ and (c) $n_{ch} \geq 20$. The absolute difference between the two distributions as a function of $\Delta\eta$ is shown in the bottom panels.

A comparison between the corrected and generated pseudorapidity distributions is shown in figure 7.5 for the two centre-of-mass energies and the different phase-space regions explored at 7 TeV. In all cases, the red histogram corresponds to the true observable and the black markers to the corrected distribution after applying the procedure described in section 7.1. At the bottom of each plot, the absolute difference between the two distributions, $(truth - corrected)$, in each $\Delta\eta$ bin is presented.

At both 900 GeV and 7 TeV with $n_{ch} \geq 2$ (figures 7.5a and 7.5b) there is a very good agreement between the corrected and truth distributions, with an absolute difference in most of the bins of ± 0.05. This is not the case, however, for the high multiplicity phase-space (figure 7.5c), where there is good closure, of the order of 0.05, in the region of the tails ($|\Delta\eta| > 1.0$), but more significant discrepancies around the central peak, with a maximum absolute difference in the two central bins of 0.2. The origin of this non-closure is described below, in section 7.2.2.

As the method does not provide a perfect closure, the remaining differences are considered as a systematic uncertainty; this will be discussed in section 8.2.1 in more detail.

7.2.2 Effect of Secondary Particles

The success of the probabilistic track removal correction method is limited by how well randomly removing tracks according to their track reconstruction efficiency approximates the true effects of the detector. A simple way to test this statement is by applying a single "track" removal iteration to the truth level distribution and compare this *reduced-truth* to the full simulation. If the correction method is a good approximation to the effects of the apparatus, then these two distributions should be very similar.

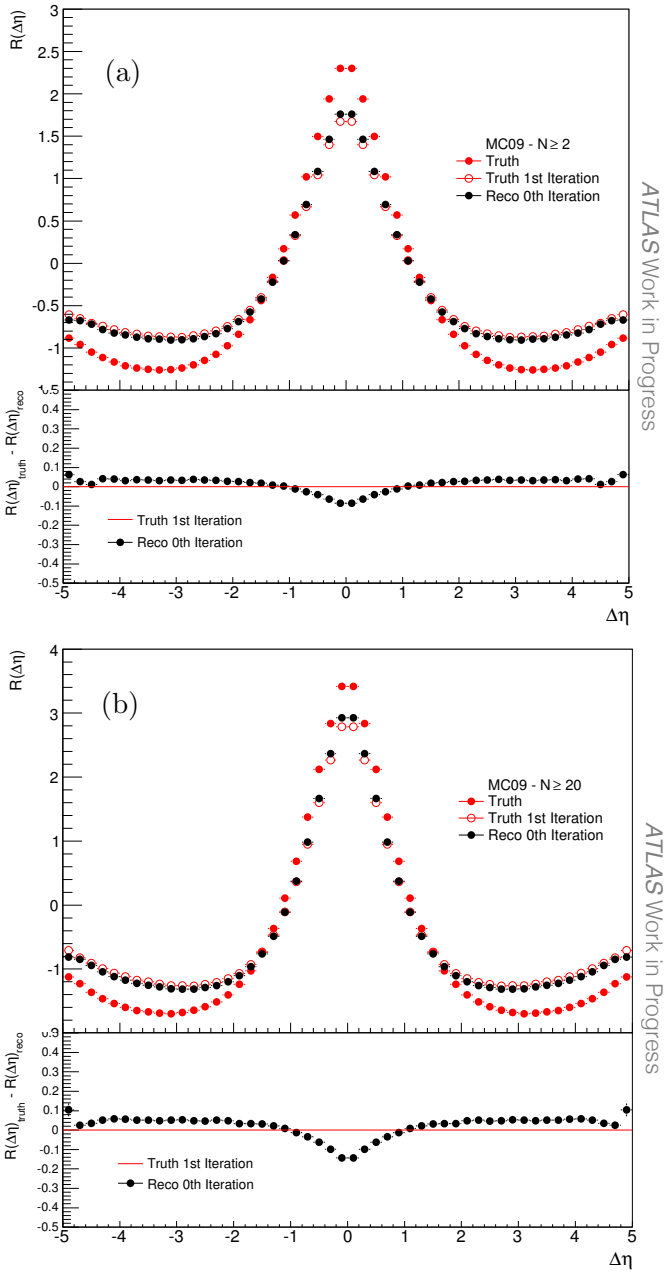


Figure 7.6: Comparison between the $R(\Delta\eta)$ obtained from full simulation (black circles) and reduced truth (open red circles) at 7 TeV, for (a) $n_{ch} \geq 2$ and (b) $n_{ch} \geq 20$. The absolute difference between reconstructed and reduced truth is shown in the bottom panels.

Figure 7.6 shows a comparison between the $R(\Delta\eta)$ obtained using the full simulation (solid black circles) and the one obtained from the reduced-truth sample, after one iteration of track removal (open red circles); the solid red

circles correspond to the original hadron-level truth distribution. As shown in the bottom panels of the plots, the reduced-truth distribution agrees reasonably well with the reconstructed one; the absolute difference is flat across the tails and reaches a maximum around the central peak (a similar behaviour to the non-closure, and, in fact, of similar magnitude), which is more pronounced in the high multiplicity case.

In both cases, the reconstructed distribution is slightly higher than the reduced-truth in the central region around the peak. This indicates that there are particles at the detector-level that are not present at the hadron-level. This is really not surprising because the probabilistic track removal method accounts for particles that are present in the truth-level but, due to inefficiencies in the reconstruction, are not observed by the detector; it does not account for particles that *are* reconstructed but do not appear in the hadron-level event definitions[1].

Secondary tracks are a source for extra particles at the detector-level, and they can be classified in two categories: physics secondaries and material secondaries. *Physics secondaries* arise from the decay of long-lived particles, such as neutral kaons. As previously discussed, a generator-level cut on the average lifetime of particles is implemented so that particles with a decay length $c\tau > 10$ mm are set stable. This means that their decay products are absent from the hadron-level event records. However, these particles are handed over to Geant4 (section 4.1.2) and are decayed by the simulation, so that if they have impact parameters of less than 1.5 mm, they will be present in the detector-level and used for analysis.

Correcting for, or removing, physics secondaries would require the ability

[1]The contribution from fake tracks was shown to be negligible in [84].

to identify the different particle species involved and would introduce model dependent factors. An estimate of the size of the effect can be made, however, using truth-level Monte Carlo. Instead of applying the $c\tau$ cut at the generator-level, a cut on the transverse and longitudinal impact parameters of 1.5 mm is used to generate a new sample where physics secondaries are retained. Using different tunes of PYTHIA6 (DW, MC09 and AMBT1), the pseudorapidity correlation function is calculated with and without the $c\tau$ cut. A comparison between these two distributions at 7 TeV, for $n_{ch} \geq 2$ and $n_{ch} \geq 20$, is presented in figure 7.7. For the different tunes, the open squares correspond to the observable calculated using the $c\tau$ cut sample, while the solid markers represent the distributions from the impact parameter cuts sample.

The absolute difference between the $c\tau$ and "raw" distributions has a similar shape to both the non-closure and the difference between reduced-truth and full simulation. However, the effect is smaller accounting for only 10% of the total non-closure in the $n_{ch} \geq 20$ case.

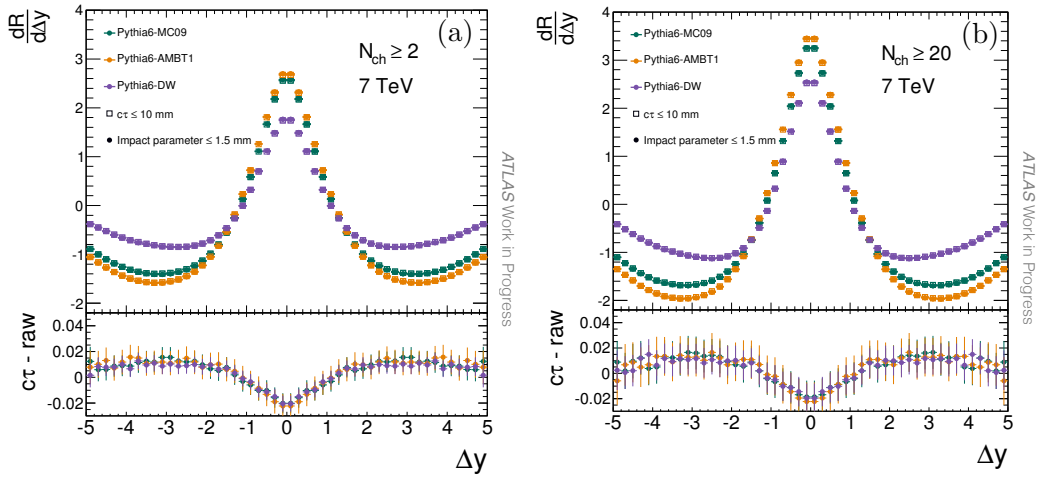


Figure 7.7: Estimate of the effect of physics secondaries on $R(\Delta\eta)$. Different Monte Carlo samples were re-generated without the $c\tau$ cut and using a cut of 1.5 mm on the impact parameters. The differences between the two distributions are shown in the bottom panels. From [93].

Material secondaries are the result of showers induced by the interaction of a primary particle with the detector material. These particles are more likely to contribute to the central peak since they are emitted close to the particle that caused the shower.

The ratio of physics secondaries to material secondaries can be determined indirectly if the total fraction of secondary particles per event is known. The Minimum Bias analysis [84] measured this total fraction as a function of p_{T} and η. Using the raw hadron-level event samples, without the $c\tau$ cut, the fraction of physics secondaries per event as a function of p_{T} and η can be estimated, using the decays of K_s, since, by far, they are the most abundant long-lived particles in the sample.

By comparing the two distributions, one can infer the fraction of material secondaries. This is illustrated in figure 7.8 for the 7 TeV, $n_{ch} \geq 2$ case. At the bottom of each distribution, the ratio of physics secondaries to all secondaries is shown; apart from at high-p_{T}, the physics secondaries account for between 20% and 50% of the total number of secondary tracks, which implies that most of the secondary particles come from material interactions.

From these studies, it is reasonable to conclude that the majority of the observed non-closure of the correction method is due to the effect of secondary particles.

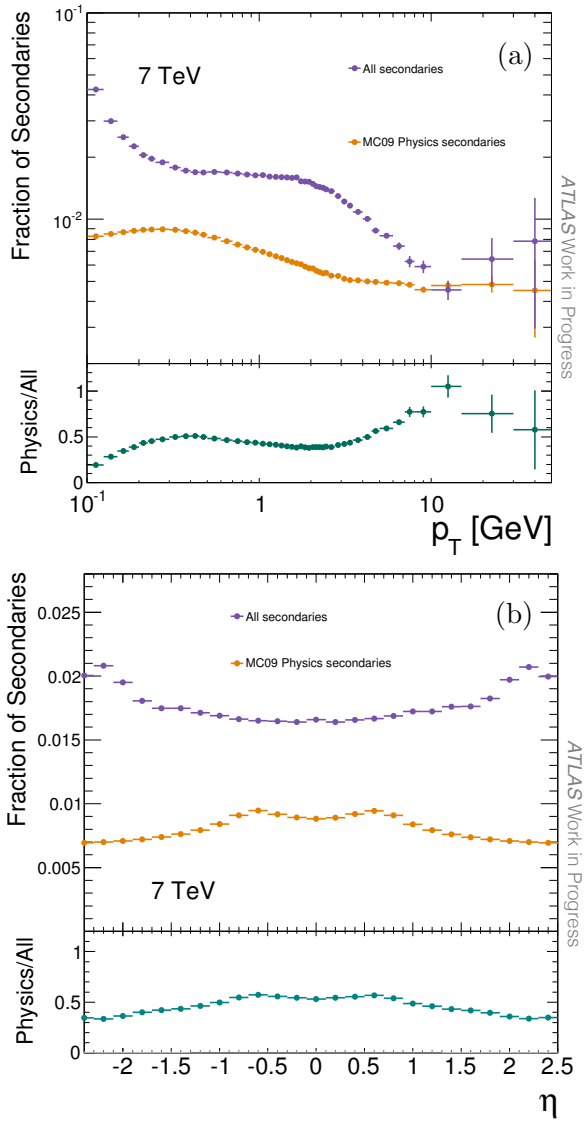


Figure 7.8: Fraction of secondaries as a function of (a) p_{T} and (b) η. The bottom panel shows the ratio of physics secondaries to the total. From [93].

Chapter 8

Uncertainties

8.1 Extrapolation to N$= -1$

The parameters a_i of the third degree polynomial fits in the probabilistic track removal correction method (section 7.1.2) are used to construct an equation $f(N)$, for each bin in the correlation distributions, of the form

$$f(N) = \sum_{i=0}^{3} a_i \mathrm{N}^i \tag{8.1}$$

where N corresponds to the iteration number. The corrected value of the bin is the result of evaluating this equation in N$= -1$. Given that the parameters of the fit are correlated, the statistical uncertainty associated to the extrapolation, σ_{-1}, is determined as [94]

$$\sigma_{-1}^2 = \mathbf{n}\mathbf{M}\mathbf{n}^{\mathrm{T}} = \sum_{i=0}^{3}\sum_{j=0}^{3} n_i M_{ij} n_j \tag{8.2}$$

where the coefficient vectors, $\mathbf{n}$, and the fit's covariance matrix, $\mathbf{M}$, are given

by

$$\mathbf{n} = \begin{pmatrix} 1 \\ \mathrm{N} \\ \mathrm{N}^2 \\ \mathrm{N}^3 \end{pmatrix} \quad , \quad \mathbf{M} = \begin{pmatrix} \sigma_0^2 & \mathrm{cov}_{01} & \mathrm{cov}_{02} & \mathrm{cov}_{03} \\ \mathrm{cov}_{10} & \sigma_1^2 & \mathrm{cov}_{12} & \mathrm{cov}_{13} \\ \mathrm{cov}_{20} & \mathrm{cov}_{21} & \sigma_2^2 & \mathrm{cov}_{23} \\ \mathrm{cov}_{30} & \mathrm{cov}_{31} & \mathrm{cov}_{32} & \sigma_3^2 \end{pmatrix} \tag{8.3}$$

with σ_i the variance of the i-th parameter, and cov_{ij} the covariance between the i-th and j-th parameters.

8.2 Systematics

In this section, the sources of systematic uncertainty identified for this analysis are presented. To illustrate the procedures followed to estimate these uncertainties, the pseudorapidity correlation function $R(\Delta\eta)$, obtained by integrating $\Delta\phi$ between 0 and π, is used. The same techniques presented here are applied to all the projections of the correlation function, as well as to the two-dimensional distributions at all energies and regions of phase-space.

8.2.1 Non-Closure of the Correction Method

The small discrepancies between the generator-level and the full simulation distributions after applying the corrections described in section 7.1, are considered as a systematic uncertainty to express the level of confidence on the correction method.

The absolute difference between the truth and corrected distributions is used to estimate the amount of non-closure. The reason why the difference is used, instead of a fractional uncertainty, is that for some $\Delta\eta$ bins $R(\Delta\eta)$ is very close to zero, which would result in a spuriously large non-closure for that bin. The issue is that data and Monte Carlo do not agree on the

values of $\Delta\eta$ for which $R(\Delta\eta)$ is close to zero. This means that estimating the data uncertainty from Monte Carlo from figure 7.5 in section 7.2.1 would not be adequate. Instead, the difference as a function of the corrected values of R is used. Then, for a given value of $R(\Delta\eta)_{corrected}$ in data, the systematic uncertainty associated to the non-closure, i.e. the magnitude of the difference between truth and corrected MC, can be read from this distribution.

For this particular $\Delta\eta$ projection it is observed that for the $n_{ch} \geq 2$ phase-space, at both 900 GeV (figure 8.1a) and 7 TeV (figure 8.1b), the absolute amount of non-closure is not strongly dependent on the value of $R(\Delta\eta)_{corrected}$ and generally is in the region of or less than 0.05, thus, a flat systematic uncertainty of 0.05 is assigned to all bins in data. In the case of $n_{ch} \geq 20$ (figure 8.1c) a flat uncertainty of 0.05 can be used for $R(\Delta\eta)_{corrected} < 0$, but as $R(\Delta\eta)_{corrected}$ becomes positive there is a dependence of the non-closure that can be parameterised with a linear equation. The complete set of plots for all $\Delta\eta$ and $\Delta\phi$ projections of the correlation function at 7 TeV and $n_{ch} \geq 2$ is given in Appendix B.

Model Dependence

To explore any possible model dependence on the amount of non-closure, the difference as a function of $R(\Delta\eta)_{corrected}$ was determined for different Monte Carlo models and tunes. As shown in figure 8.2 for $n_{ch} \geq 2$ at 7 TeV, all of the models exhibit the same behaviour and, for this particular example, lie within the band of 0.05 derived with the MC09 tune.

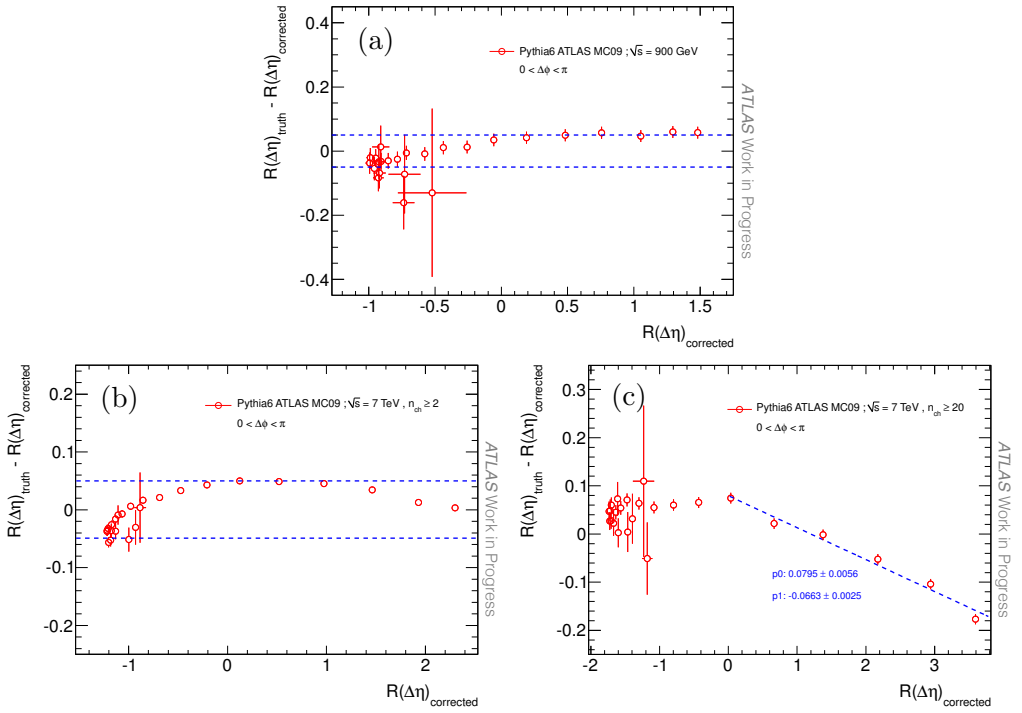


Figure 8.1: Difference between the truth and corrected pseudorapidity distributions as a function of $R(\Delta\eta)_{corr}$, for the PYTHIA6 MC09 tune at (a) 900 GeV and 7 TeV for (b) $n_{ch} \geq 2$ and (c) $n_{ch} \geq 20$.

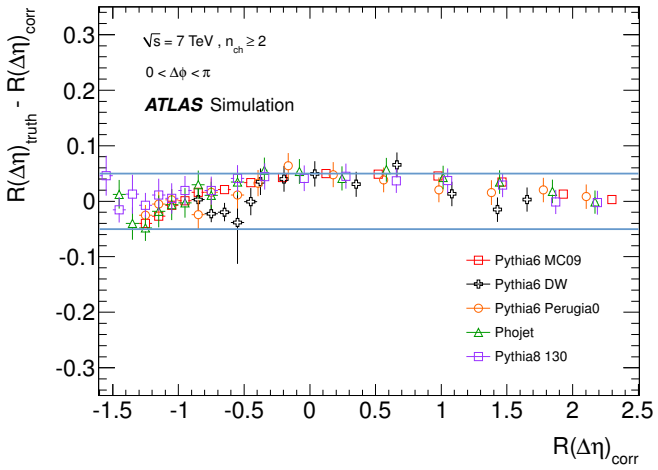


Figure 8.2: Difference between the truth and corrected pseudorapidity distributions as a function of $R(\Delta\eta)_{corr}$ for different Monte Carlo tunes at 7 TeV for $n_{ch} \geq 2$. Bins with similar values of $R(\Delta\eta)_{corr}$ have been merged to reduce the statistical fluctuations and better illustrate the behaviour of the models.

8.2.2 Uncertainties on the Efficiencies

To evaluate the systematic uncertainty associated to uncertainties in the trigger, vertex and track reconstruction efficiencies *from data*, the two-particle correlation function is computed varying each of these quantities, one at a time, according to their uncertainties. These uncertainties are considered to be the same for both energies. A detailed description can be found in [84].

The systematic uncertainties associated to the track reconstruction efficiency, summarised in table 8.1, come from different sources. To obtain the total tracking uncertainty in each (p_{T}, η) bin, all of these uncertainties are added in quadrature. In the case of the trigger and vertex reconstruction efficiencies, the systematic uncertainties are of the order of 1% for events with a track multiplicity of 2 and rapidly decrease as the multiplicity increases.

Source	Uncertainty
Material description in MC	increases at high-η and decreases at high-p_{T} max. 15% in the range $2.3 < \lvert\eta\rvert < 2.5$ and $100 < p_{\mathrm{T}} < 150$ MeV
Track Selection	1% in all (p_{T}, η) bins
p_{T} resolution in MC	5% in the first p_{T} bin: $100 < p_{\mathrm{T}} < 150$ MeV
High-p_{T} tracks ($p_{\mathrm{T}} > 10$ GeV)	10% due to the track-fit χ^2 probability cut; η and p_{T} dependent uncertainties due to mis-measured tracks

Table 8.1: Uncertainties on the track reconstruction efficiency.

The upper panel of figure 8.3 shows a comparison between the nominal 7 TeV $R(\Delta\eta)$ distribution (when $n_{ch} \geq 2$) and the ones obtained by increasing and decreasing the track reconstruction efficiency, used in the track removal iterations, by its systematic uncertainty. The differences between the distributions, shown in the bottom panel, constitute the systematic uncertainty associated to tracking. For the majority of the bins, the tracking systematic is the dominant uncertainty of the two-particle correlations measurement.

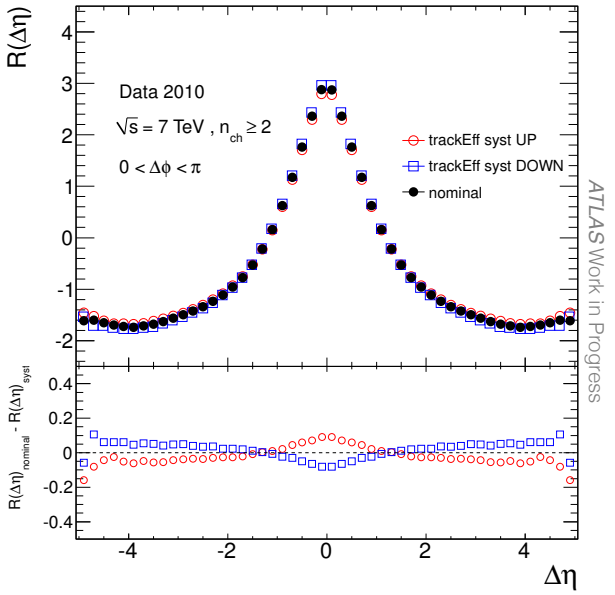


Figure 8.3: Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the tracking efficiency by its systematic uncertainty.

For the calculation of the systematic uncertainty associated to the track reconstruction in figure 8.3, the tracking uncertainties were assumed to be correlated across all η regions, which implies that the track reconstruction efficiency was shifted in the same direction in each pseudorapidity bin. However, it could be the case that due to, for example, differences in the amount of material or mis-alignment, different regions of the detector could behave distinctly. In figure 8.4, the track reconstruction efficiency was shifted in opposite directions for the barrel ($|\eta| < 2.0$) and end-caps ($2.0 < |\eta| < 2.5$) and compared to the nominal $R(\Delta\eta)$ distribution. The observed differences are smaller than those in figure 8.3 so this effect can be considered to be covered by the existing systematic uncertainty.

The distributions obtained by varying the vertex and trigger reconstruction efficiencies are shown in figures 8.5 and 8.6, respectively. As with the tracking systematics, the total uncertainty in each bin corresponds to the differences between the nominal and systematically shifted pseudorapidity distributions.

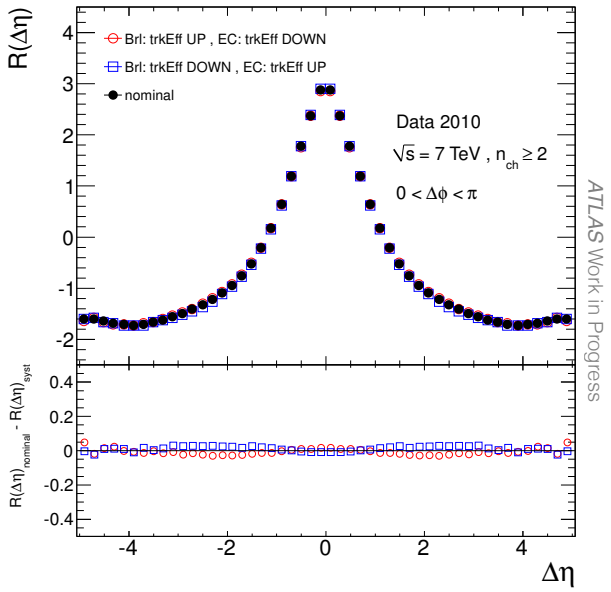


Figure 8.4: Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the tracking efficiency by its systematic uncertainty in opposite directions for the barrel and end-cap regions of the detector.

The size of these uncertainties is quite small, becoming negligible for the $n_{ch} \geq 20$ phase-space.

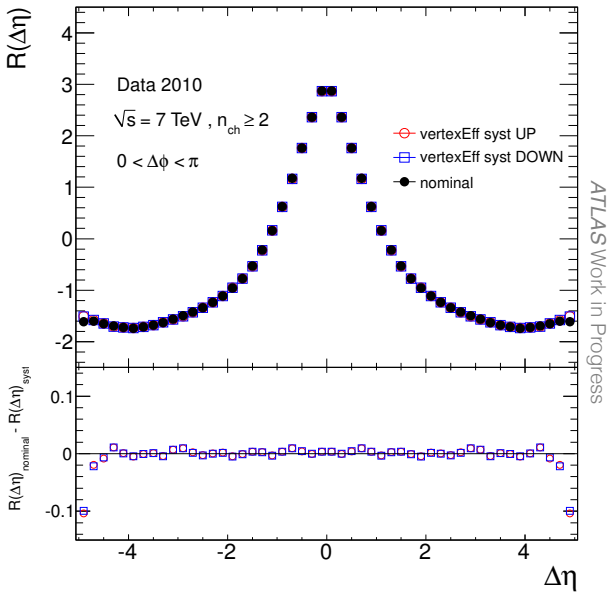


Figure 8.5: Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the vertex reconstruction efficiency by its systematic uncertainty.

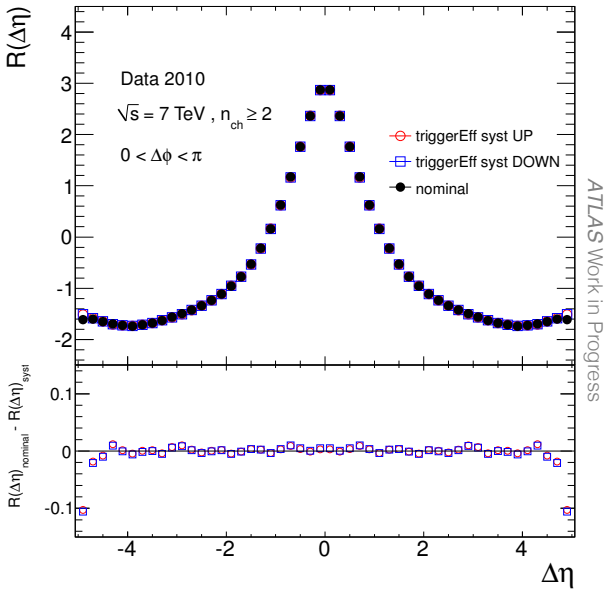


Figure 8.6: Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the trigger reconstruction efficiency by its systematic uncertainty.

8.2.3 Additional Checks

Studies on other possible sources of uncertainty are described in this section. First, the effect of the choice of the number of iterations and fitting function used in the probabilistic track removal correction method is explored. Then, different random number generators are used to remove tracks to test if there is any dependence due to the large number of tracks in the samples. Finally, the effect of possible correlated track reconstruction efficiencies is discussed.

Figure 8.7 illustrates the closure obtained, using the MC09 at 7 TeV sample, when applying five and four track-removal iterations to correct the pseudorapidity distributions (the nominal analysis uses six iterations in total). In both cases, there is good agreement and the difference between the truth and corrected distributions lies within the 0.05 uncertainty band previously obtained (section 8.2.1). This suggests that this effect can be considered to be already covered by the systematic uncertainty associated to the non-closure.

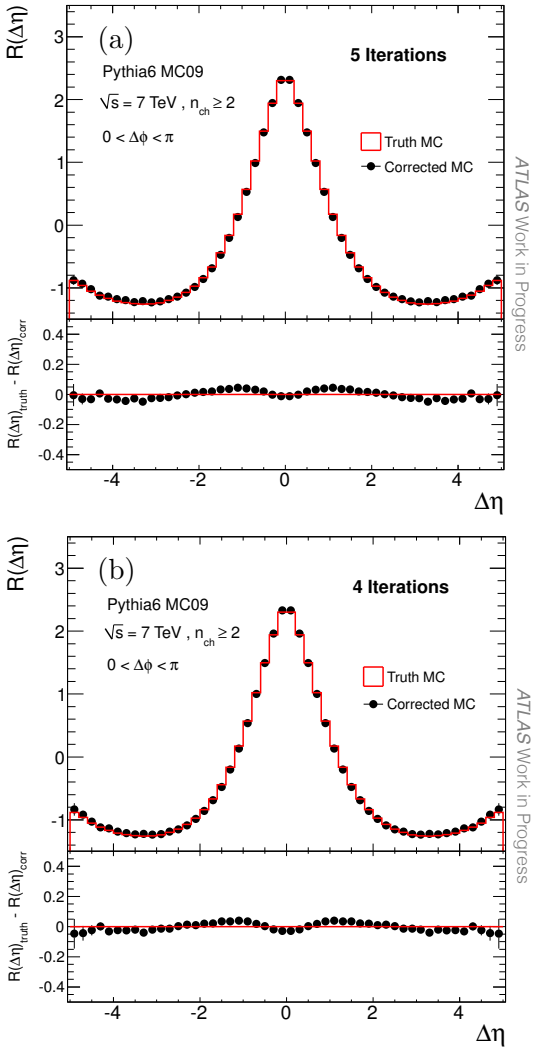


Figure 8.7: Closure tests for the $R(\Delta\eta)$ distribution when using (a) five and (b) four track-removal iterations. In both cases, the difference between truth and corrected is of or less than 0.05.

As shown in figure 8.8, when using four or five iterations to correct the data, the resulting distributions are within the uncertainties in each bin (the green bands correspond to the total uncertainty in each bin of the nominal analysis, obtained by adding in quadrature the contributions from the statistic - section 8.1 - and systematic - sections 8.2.1 and 8.2.2 - uncertainties).

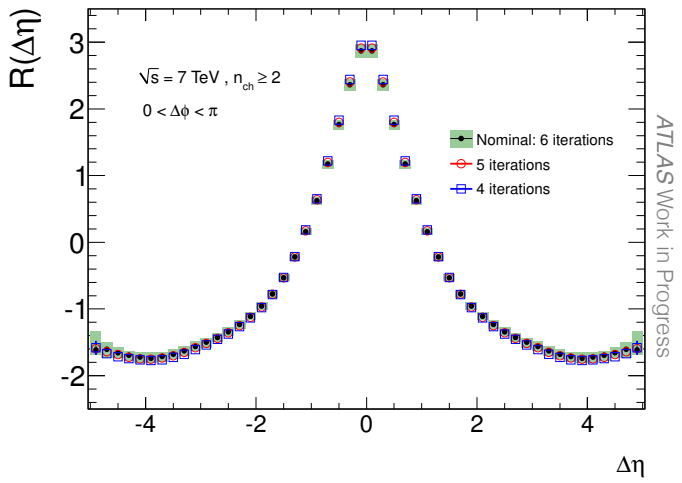


Figure 8.8: 7 TeV data $R(\Delta\eta)$ distribution corrected using four, five and six iterations of the track removal procedure.

A similar conclusion is reached for the choice of the fitting function. Polynomials of a degree smaller than three cannot describe the behaviour of the different bins in the distributions, so only the performance of higher degree polynomials is evaluated. In figure 8.9, the effect of using a fourth degree polynomial to extrapolate to the corrected values is shown for 7 TeV Monte Carlo and data. The observed closure in MC is of the order of 0.05 and the data distributions agree, within uncertainties, with each other. This effect is also considered to be covered by the non-closure systematic uncertainty.

The choice of the random number generator used for removing tracks in the correction method is studied to determine whether or not it introduces any bias on the measurement. The nominal analysis is done with the function `Rndm` inside the `TRandom` class in `ROOT` [95]. Two different classes are used to calculate the correlation function: `TRandom2` and `TRandom3`. The former is based on the Tausworthe random number generator [96] and the latter on the Mersenne Twister generator [97].

Figure 8.10 shows the shape of the $R(\Delta\eta)$ distribution when using these different random number generators for data and Monte Carlo MC09 at 7

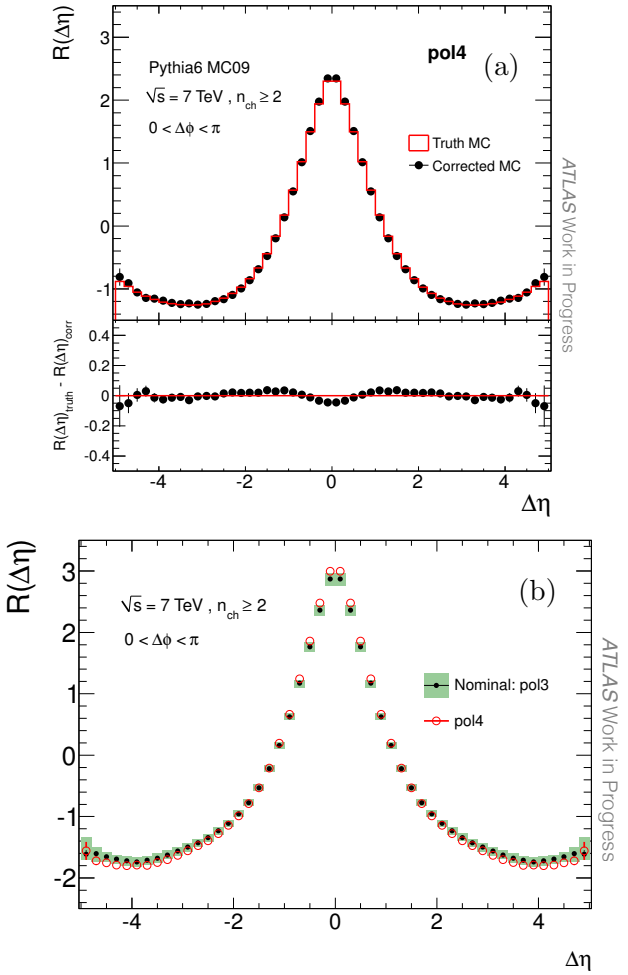


Figure 8.9: Effect of using fourth degree polynomial fits in the extrapolation to the corrected value of $R(\Delta\eta)$ at 7 TeV in (a) MC09 and (b) data.

TeV. The small variations, of the order of 0.01-0.02, are present in both the data and the simulation. If the choice of the random number generator, or even the random number seed, makes a contribution to the small non-closure in Monte Carlo, it will make the same contribution to the data. As such, any systematic effect due to the randomness of the correction method must be already accounted for in the non-closure systematic uncertainty; if the correction method closed perfectly, then there could not be any contribution from the random number generator.

As an additional check of the independence of the measurement on the

random number generator, figure 8.11 illustrates the closure tests for $R(\Delta\eta)$ using the `TRandom2` and `TRandom3` classes to remove tracks according to their tracking efficiency. Regardless of the choice of the generator, there is good agreement between the distributions and, furthermore, the amount of non-closure is consistent with what is obtained in the nominal analysis.

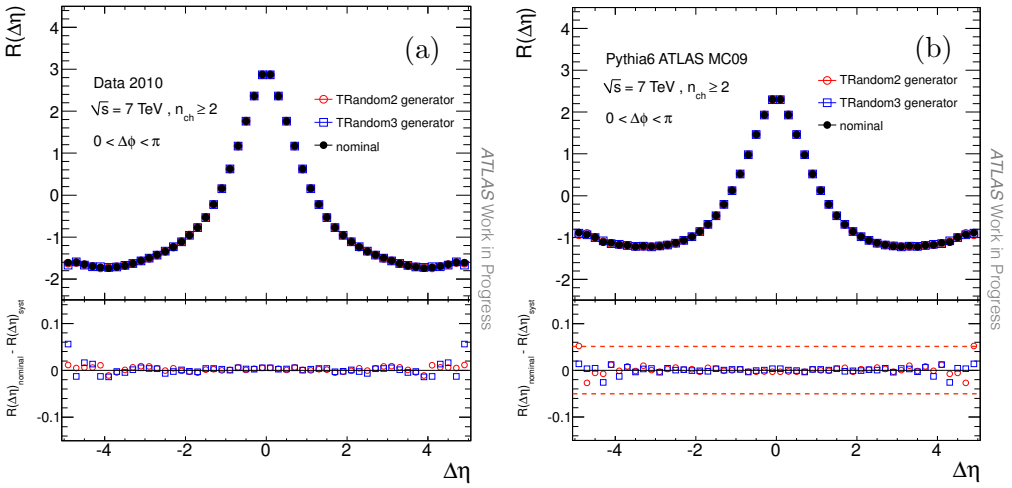


Figure 8.10: Pseudorapidity correlation functions, at 7 TeV ($n_{ch} \geq 2$), obtained by varying the random number generator used in the probabilistic track removal iterations in (a) data and (b) the Monte Carlo tune MC09.

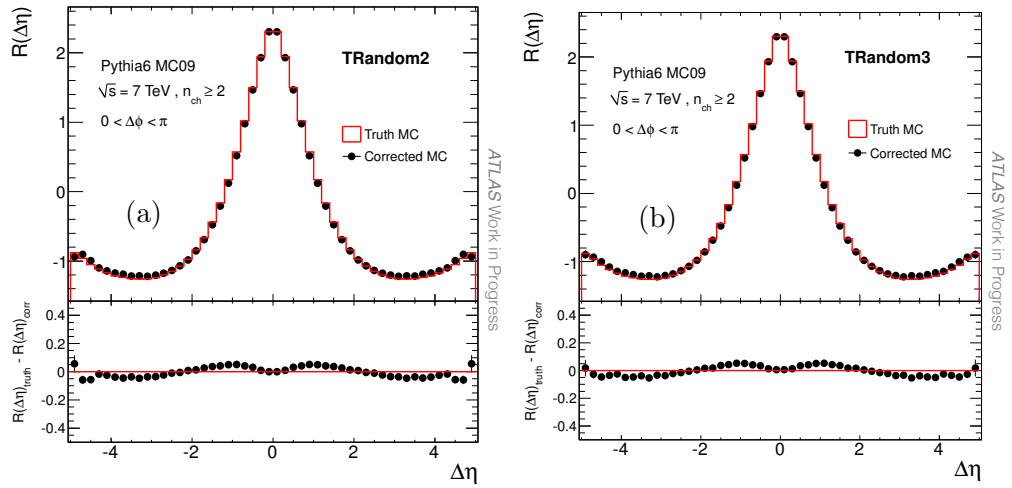


Figure 8.11: Generated pseudorapidity correlation function compared to the ones corrected using the (a) TRandom2 and (b) TRandom3 random number generators to throw away tracks.

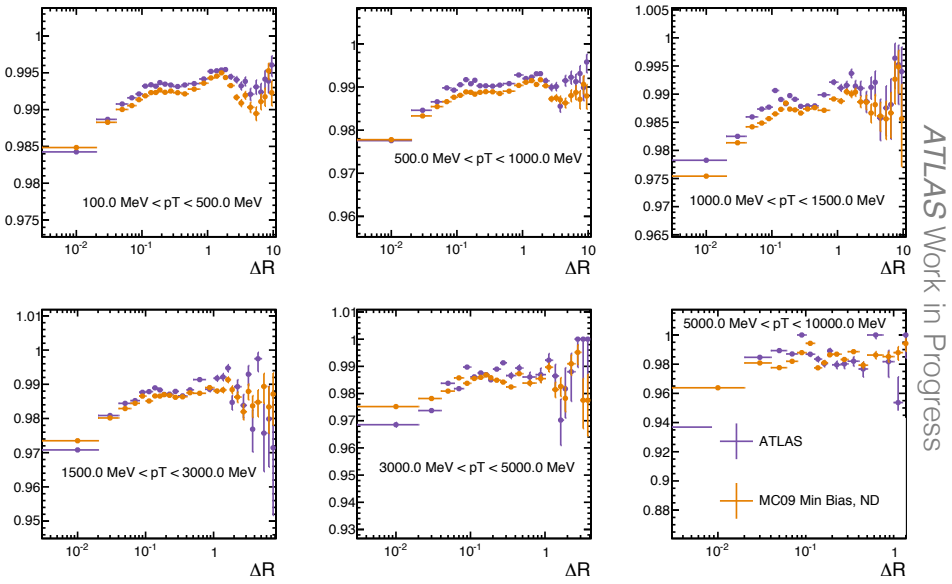


Figure 8.12: Fraction of nearest neighbour tracks with a hit on the Pixel b-layer as a function of the $\Delta\mathcal{R}$ separation ($\Delta\mathcal{R}^2 = \Delta\eta^2 + \Delta\phi^2$) for 7 TeV, $n_{ch} \geq 2$ data and non-diffractive Monte Carlo MC09 in different p_{T} bins. From [93].

Finally, the possibility of having correlated track reconstruction inefficiencies is explored. The track removal correction method assumes that the track losses are uncorrelated, which raises the question: if, for example, two tracks are very close together in (η, ϕ), does the probability of both of them getting lost increase? To study this, the correlation in the Pixel b-layer hit requirement between nearest tracks is investigated. The same event and track selection criteria as in section 6.2 are applied, with the exception of requiring a b-layer hit. Then, if a track *expected* a b-layer hit, the nearest neighbour track, i.e. the track with the smallest $\Delta\mathcal{R}$ (where $\Delta\mathcal{R}^2 = \Delta\eta^2 + \Delta\phi^2$), is identified from the remaining tracks. A histogram is filled with these $\Delta\mathcal{R}$ values. A second histogram is created with only the $\Delta\mathcal{R}$ for the track pairs in which the neighbour *did* have a b-layer hit. The ratio of these two histograms is plotted in figure 8.12 for 7 TeV data and non-diffractive MC09. It represents the fraction of nearest neighbour tracks with a b-layer hit as a function of

$\Delta\mathcal{R}$, in different p_{T} bins. It can be seen that, if a particle expected a b-layer hit (no matter if it actually had it or not), above 90% of the times the nearest neighbour track did have it, and this fraction does not vary much with $\Delta\mathcal{R}$. This suggests that losses of nearby tracks are not correlated. Furthermore, data and Monte Carlo exhibit the same behaviour and the small differences between them, of the order of 0.1%, are the same across all $\Delta\mathcal{R}$, which rules out any issues with the detector simulation, but rather implies that they arise from different p_{T} distributions.

From these studies, it is concluded that the only systematic contributions to the total uncertainty, considered in the two-particle correlations measurement, are the non-closure of the correction method and the uncertainties on the reconstruction efficiencies.

Chapter 9

Results

9.1 Two-Dimensional Correlation Functions

The corrected p_{T} inclusive two-particle angular correlation functions for data and the PYTHIA6 MC09 Monte Carlo tune are shown in figure 9.1 for $\sqrt{s} = 900$ GeV and 7 TeV. The total uncertainties in each bin of these two-dimensional distributions are calculated following the procedures described in chapter 8 and are presented, for reference, in appendix C.

To construct these two-dimensional distributions, the absolute values of the $\Delta\eta$ ($0 < |\Delta\eta| < 5$) and $\Delta\phi$ ($0 < |\Delta\phi| < \pi$) separations between the particle pairs are used to fill one quadrant of the foreground and background ($\Delta\eta$,$\Delta\phi$) histograms, while the other three quadrants are filled by reflection. As a consequence, the correlation function is symmetric around (0,0) by construction. To be able to properly see the complete structure in the $\Delta\phi$ axis, instead of it being split at $\pm\pi$, the distributions are shifted and plotted in the range $-\frac{\pi}{2} < \Delta\phi < \frac{3\pi}{2}$.

At both energies a complex structure is observed across the full $\Delta\eta$ and $\Delta\phi$ range, although, due to limited statistics, the 900 GeV distributions look much more noisy. Several components can be identified that reflect the contributions

from different underlying processes to the correlation structure. The feature that immediately stands out is a sharp peak around (0,0). These *near-side correlations* can be attributed to particles originating from the same high-p_{T} process. Such closely correlated particles indicate the emergence of jet-like structures within the minimum bias events. A second component are the *away-side correlations*, seen as a ridge extending across the whole $\Delta\eta$ range near $\Delta\phi \sim \pi$. This activity arises from the recoil of one parton against another. Finally, a broad Gaussian-like structure spanning the whole $\Delta\phi$ axis, centred at $\Delta\eta = 0$ with a width of approximately two units in $\Delta\eta$, is observed. These are called *short-range correlations* in $\Delta\eta$ that can be related to low-p_{T} processes such as the decay of resonances, clusters or string fragmentation.

In figure 9.1, the data distributions are shown on the left-hand side and the simulation distributions on the right. These plots are symmetric around $\Delta\eta = 0$ by construction.

As the centre-of-mass energy increases so does the height of the central peak and it becomes even more pronounced for the sample with events with a higher charged particle multiplicity (figure 9.1e). Conversely, the height of the away-side ridge is constant regardless of the collision energy or multiplicity. The MC09 tune exhibits similar structures as those observed in data, however, for the 900 GeV and 7 TeV $n_{ch} \geq 2$ samples, it does not reproduce the strength of the correlations. For the higher multiplicity sample, the simulation reproduces the height of the central peak but, as in the other two cases, differs in the shape and strength of the away-side ridge.

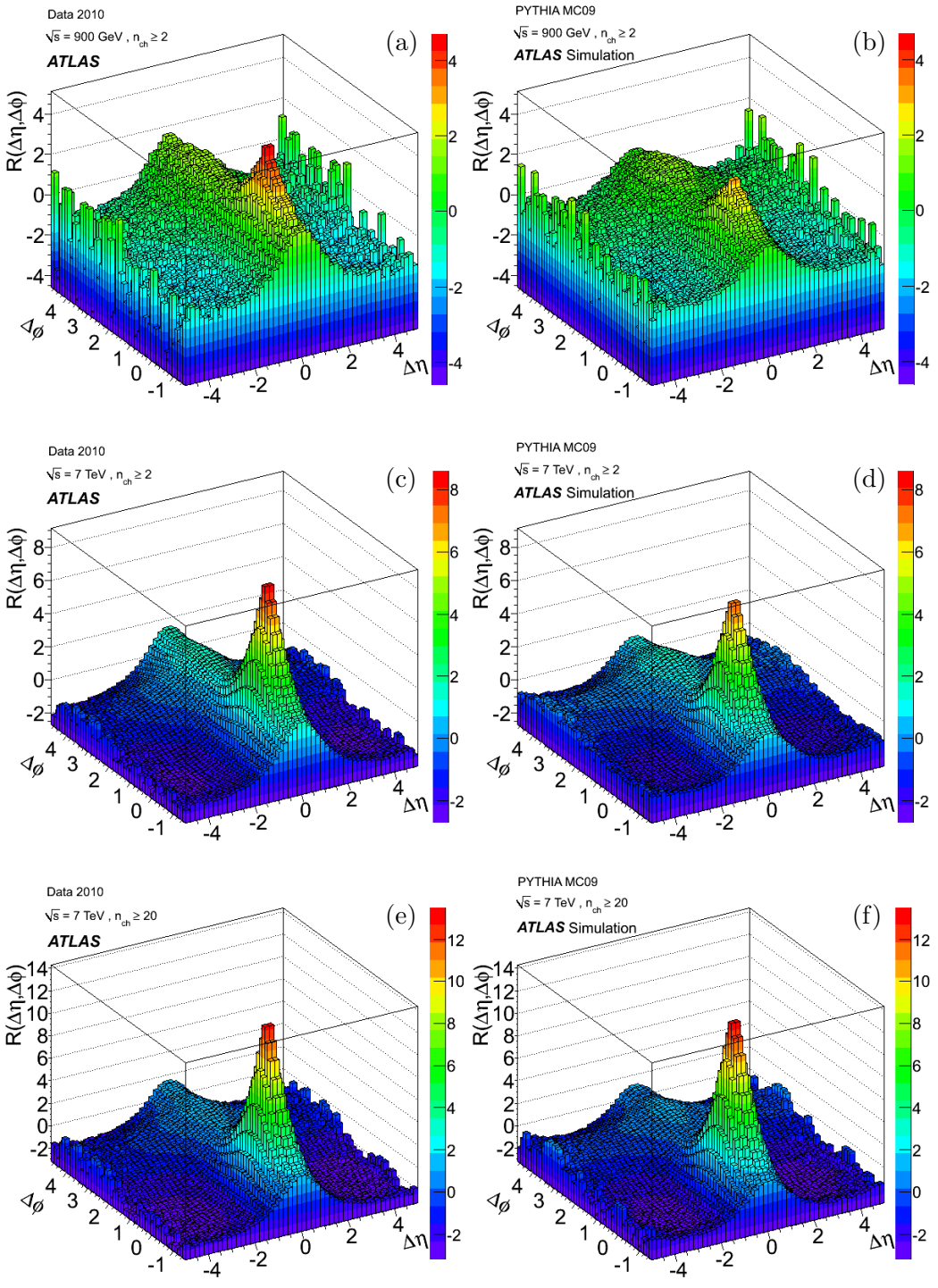


Figure 9.1: Corrected $R(\Delta\eta, \Delta\phi)$ distributions for 900 GeV and 7 TeV data (left column) and the MC09 tune (right column). These plots are symmetric around $\Delta\eta = 0$ by construction.

9.2 Projections in $\Delta\eta$ and $\Delta\phi$

In this section, different projections of the measured two-dimensional correlation function are explored and compared to the Monte Carlo tunes introduced in section 1.4.3. A comparison of the corrected data to older tunes was documented in the ATLAS note [98], written in collaboration with Dr James Monk.

Figure 9.2 illustrates the different regions into which the $\Delta\eta\Delta\phi$-space is divided in order to examine in detail the structure of the correlation function.

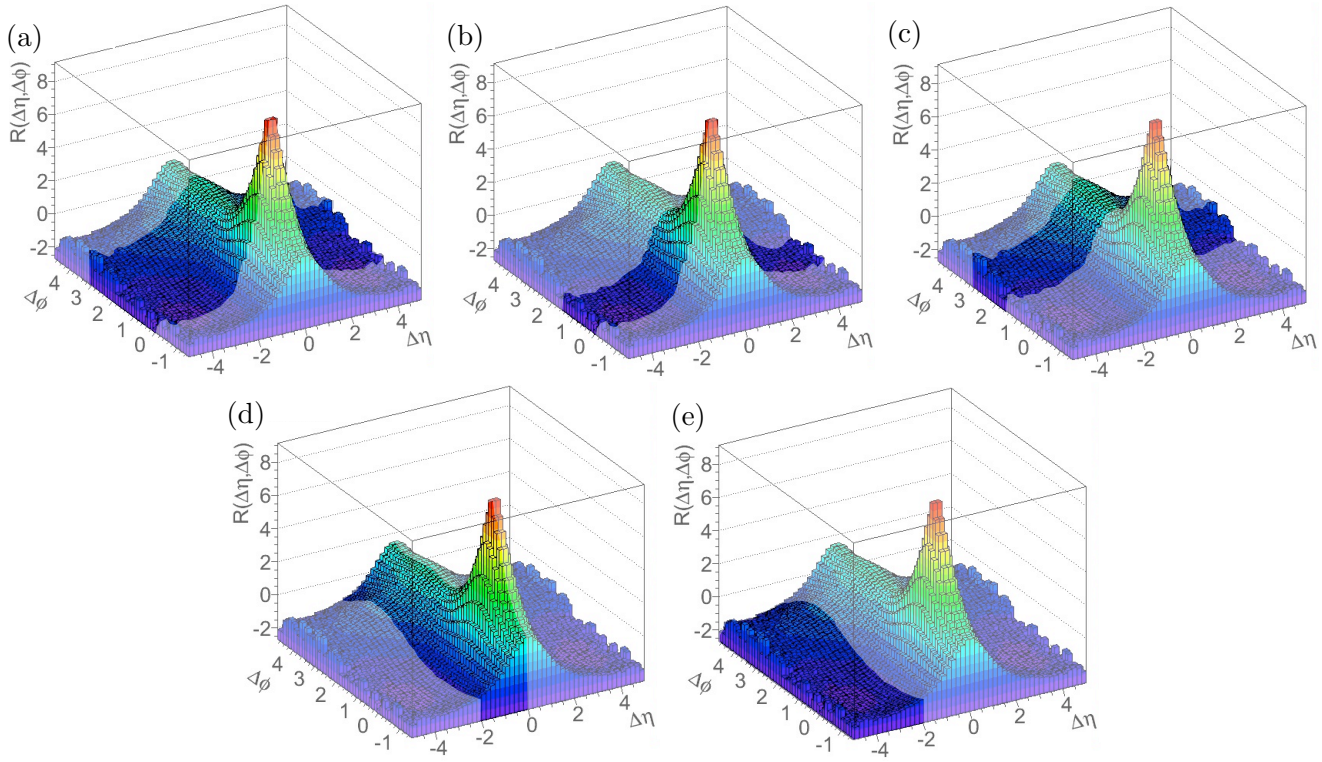


Figure 9.2: Regions used for the one-dimensional projections of the correlation function. For projections into $\Delta\eta$, $\Delta\phi$ is integrated between (a) 0 and π, (b) 0 and $\pi/2$, and (c) $\pi/2$ and π. For projections into $\Delta\phi$, $\Delta\eta$ is integrated between (d) 0 and 2, and (e) 2 and 5.

In all of the following figures, the solid black markers correspond to the corrected data values, the black error bars are statistical only (coming from the extrapolation to N$= -1$) and the solid green bands correspond to the total

uncertainty in the bin, which is obtained by adding in quadrature the statistical and systematic uncertainties described in chapter 8. The MC curves are overlaid as differently coloured histograms, each with different line styles. To quantify the discrepancies between data and the models, the absolute difference, $(Data - MC)$, is presented underneath each distribution. In general, none of the models reproduce the strength of the correlations seen in the data. Some approximate the shape of the distributions more closely than others, but for all cases HERWIG++ is the most discrepant tune with the data.

$R(\Delta\eta)$ with $0 < \Delta\phi < \pi$ (figure 9.2a)

Figures 9.3 to 9.5 show the pseudorapidity correlation function $R(\Delta\eta)$, obtained by integrating over the full $\Delta\phi$ range, for the different energies and charged-particle multiplicities. This "$\Delta\phi$-inclusive" projection has been used by other experiments to explore the structure of two-particle correlations and interpret it in terms of cluster emission models (figures 2.7 and 2.9 in section 2.2). A discussion of the validity of these models in our data is deferred to section 9.3.

As illustrated here, the peak at $\Delta\eta = 0$ becomes more pronounced as both the energy and particle multiplicity of the events increase, and it is precisely in this region where the MC models differ more from data. At 900 GeV (figure 9.3) and 7 TeV when $n_{ch} \geq 2$ (figure 9.4), the AMBT2B and Perugia 2011 tunes of PYTHIA6, together with tune 4C of PYTHIA8, are the closest to the data distributions in the central peak, while the tails are better described by tune 4C. For the high-multiplicity sample at 7 TeV (figure 9.5), tune AMBT2B exhibits a better agreement with data over the entire $\Delta\eta$ range. Of course, these "agreements" cannot be considered good on an absolute scale.

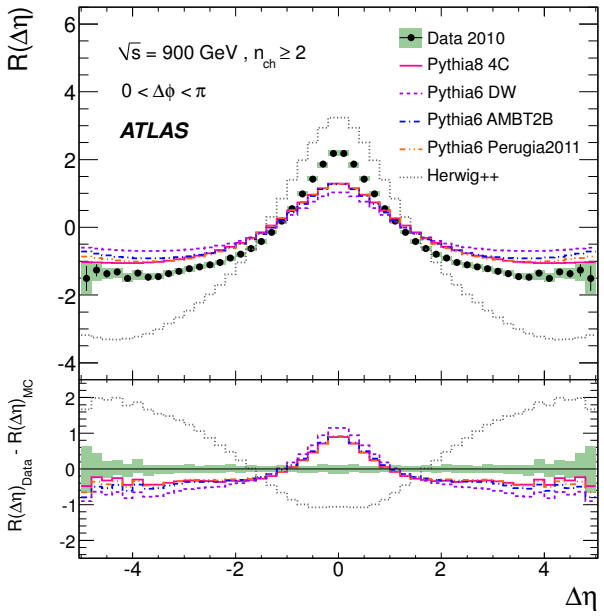


Figure 9.3: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV.

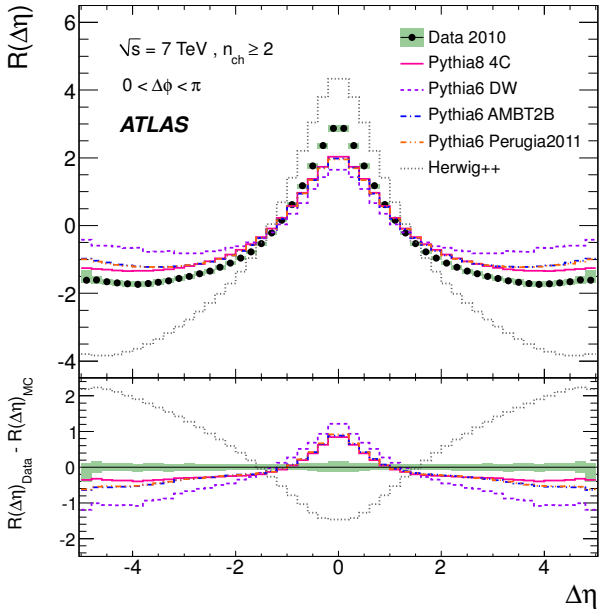


Figure 9.4: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$.

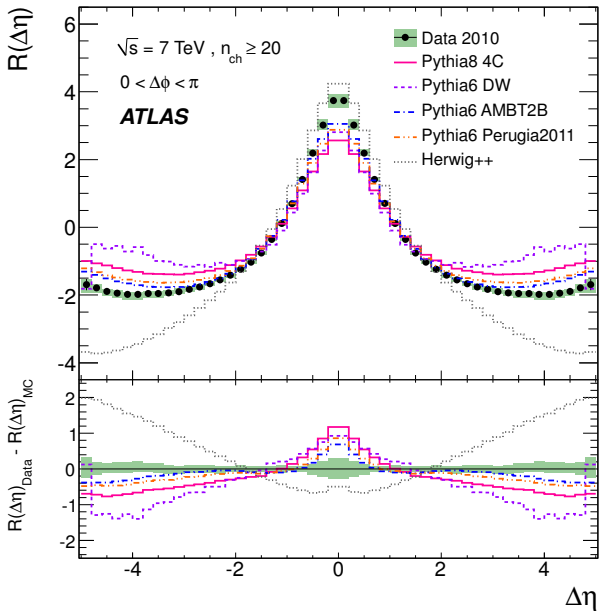


Figure 9.5: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$.

$R(\Delta\eta)$ with $0 < \Delta\phi < \frac{\pi}{2}$ (figure 9.2b)

The $\Delta\eta$ projections presented in figures 9.6 to 9.8 are obtained by integrating $\Delta\phi$ between 0 and $\frac{\pi}{2}$ (near-side). By focusing on the central peak, the distributions are narrower, higher and more pronounced, which indicates a stronger correlation between nearby particles. The amount of activity is energy dependent.

For $n_{ch} \geq 2$ (figures 9.6 and 9.7), the 4C PYTHIA8 tune does a better approximation to the data. The PYTHIA6 tunes are the ones that do a better job in the $n_{ch} \geq 20$ case (figure 9.8): DW and AMBT2B in the central peak and AMBT2B and Perugia 2011 in the tails. Also for this high-multiplicity region, HERWIG++ closes in to the data around $\Delta\eta = 0$ but the tails continue to fall.

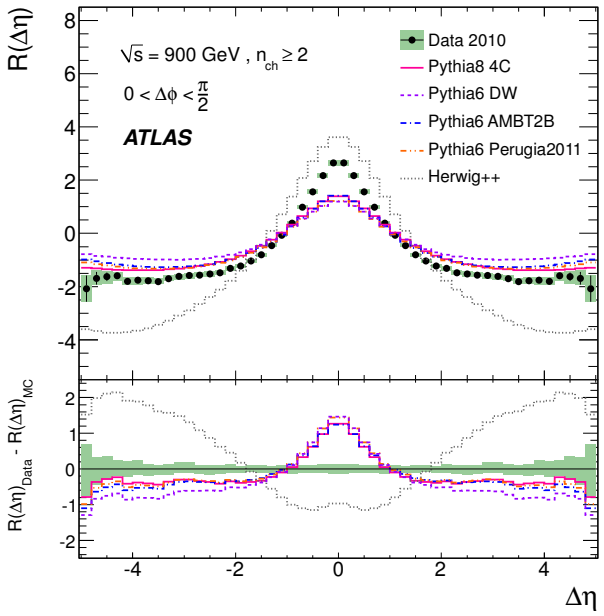


Figure 9.6: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV.

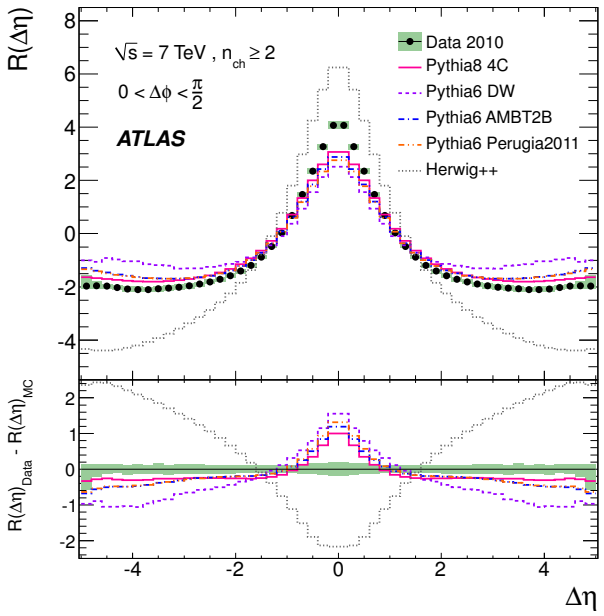


Figure 9.7: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$.

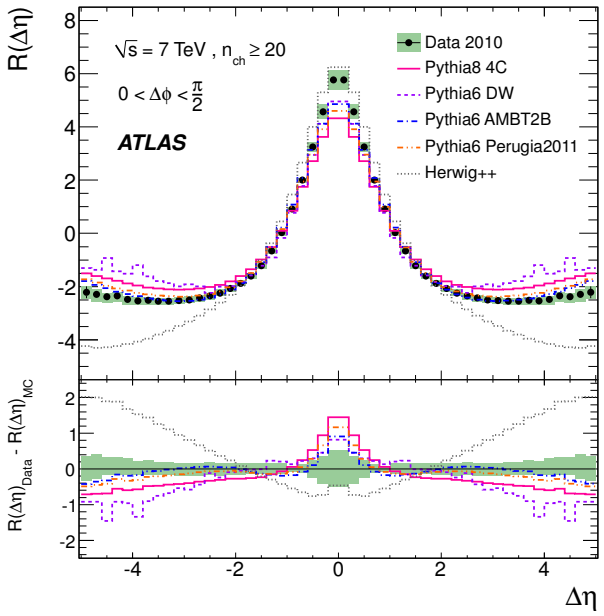


Figure 9.8: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between 0 and $\frac{\pi}{2}$, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$.

$R(\Delta\eta)$ with $\frac{\pi}{2} < \Delta\phi < \pi$ (figure 9.2c)

Contributions to the pseudorapidity correlation function from the away-side, calculated by integrating $\Delta\phi$ between $\frac{\pi}{2}$ and π, are shown in figures 9.9 to 9.11. The height and width of the distributions stay constant with energy and charged-particle multiplicity.

Most of the tunes, with the exception of HERWIG++, display a better agreement with data in this projection. The Perugia 2011 tune in particular does a reasonable description across the complete $\Delta\eta$ range for the $n_{ch} \geq 20$ 7 TeV sample (figure 9.11) and in the central peak region of events for which $n_{ch} \geq 2$ (figures 9.9 and 9.10). The DW tune predicts very little activity in this away-side region.

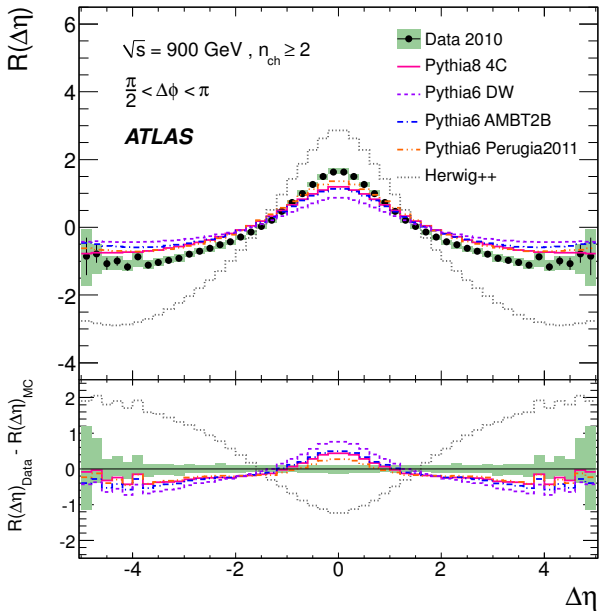


Figure 9.9: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV.

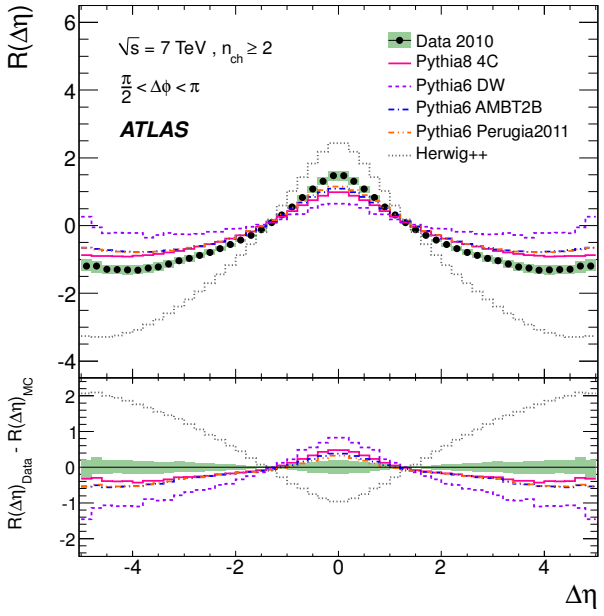


Figure 9.10: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$.

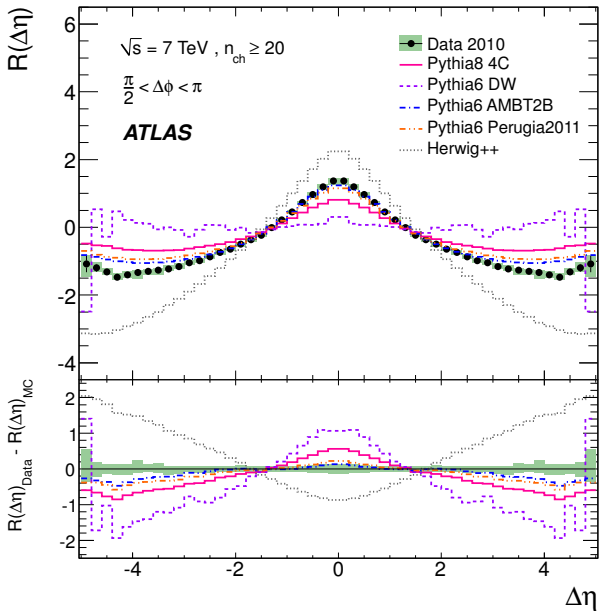


Figure 9.11: Corrected $R(\Delta\eta)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\phi$ between $\frac{\pi}{2}$ and π, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$.

$R(\Delta\phi)$ with $0 < \Delta\eta < 2$ (figure 9.2d)

The short-range correlation function, obtained by integrating $\Delta\eta$ between 0 and 2, is shown in figures 9.12 to 9.14 for the different phase-space regions. These distributions contain two peaks, one at $\Delta\phi = 0$ which becomes more pronounced if either the collision energy or the particle multiplicity is increased, and the second one at $\Delta\phi = \pi$ that is approximately constant.

At 900 GeV (figure 9.12) the difference between data and most of the tunes is flat in the region $-\frac{\pi}{2} < \Delta\phi < \frac{\pi}{2}$, and there is good agreement (within uncertainties) with data on the near-side. At 7 TeV with $n_{ch} \geq 2$ (figure 9.13), the PYTHIA8 tune 4C does a better description of the regions near the two peaks but diverges from data between them. For $n_{ch} \geq 20$ (figure 9.14), the Perugia 2011 and AMBT2B tunes agree quite well with data above $\Delta\phi \approx 1$.

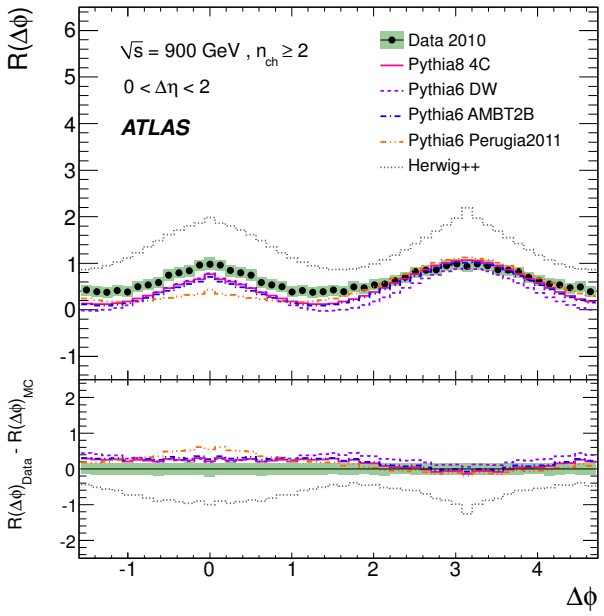


Figure 9.12: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV.

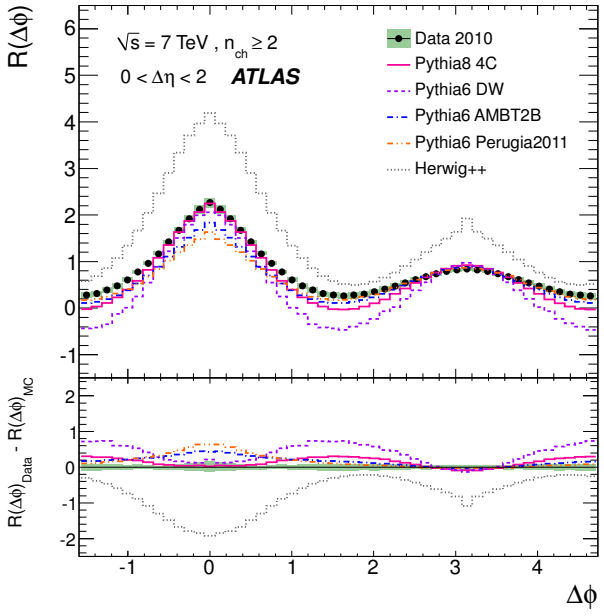


Figure 9.13: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$.

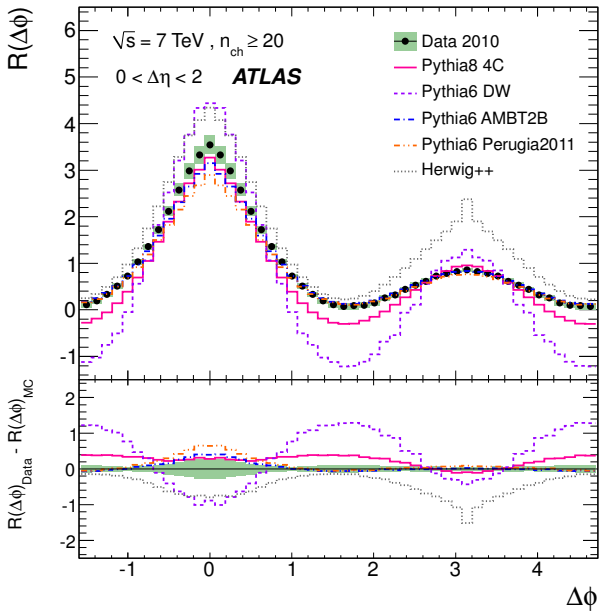


Figure 9.14: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 0 and 2, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$.

$R(\Delta\phi)$ with $2 < \Delta\eta < 5$ (figure 9.2e)

The long-range azimuthal dependence of the correlation function is determined by integrating $\Delta\eta$ between 2 and 5. The distributions in figures 9.15 to 9.17 show a trough at $\Delta\phi = 0$ while the away-side peak is still present.

For events with $n_{ch} \geq 2$ at both energies (figures 9.15 and 9.16) the differences between data and all the MC tunes are flat across the full $\Delta\phi$ range which indicates that the models are describing the shape of the correlation function correctly but not its strength. When the requirement on the particle multiplicity of the events is increased (figure 9.17), the AMBT2B and Perugia 2011 tunes of PYTHIA6 are in good agreement with data (within uncertainties) in the range $-\frac{\pi}{2} < \Delta\phi < \frac{\pi}{2}$ and the DW tune shows the largest discrepancies around the away-side peak.

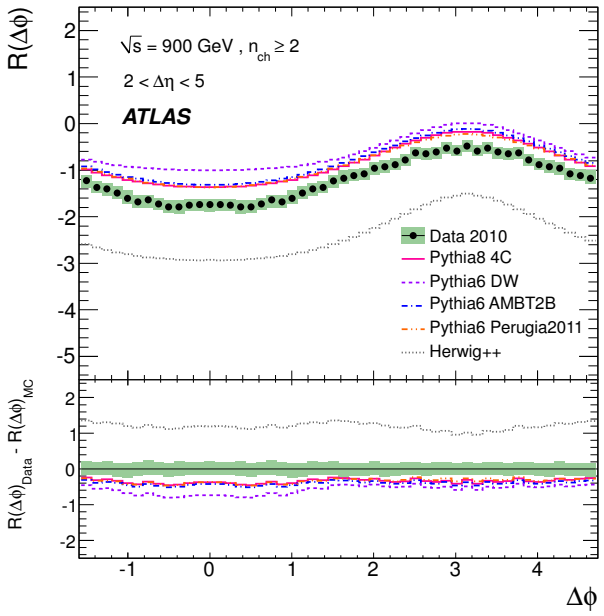


Figure 9.15: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 900$ GeV.

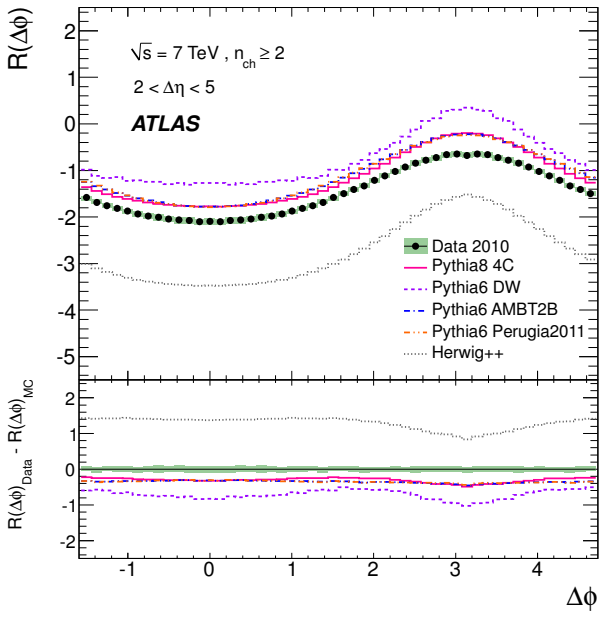


Figure 9.16: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 2$.

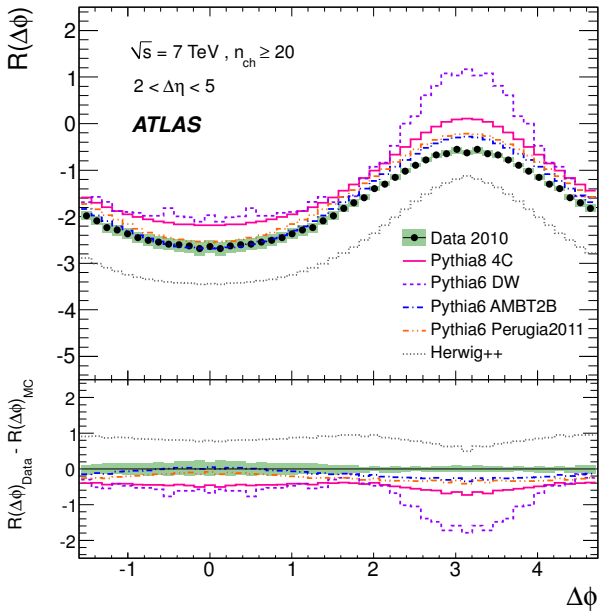


Figure 9.17: Corrected $R(\Delta\phi)$ distribution, obtained by integrating the foreground and background distributions over $\Delta\eta$ between 2 and 5, for data and the different Monte Carlo tunes at $\sqrt{s} = 7$ TeV with $n_{ch} \geq 20$.

9.3 Cluster Model Fits

The best way to compare a given model to correlations data is to provide corrected distributions and overlay Monte Carlo truth predictions for the model, as done in sections 9.1 and 9.2. However, as discussed in section 2.2, this was not always possible in the past and previous experiments opted to fit their results to phenomenological predictions. As such, the pseudorapidity two-particle correlation functions $R(\Delta\eta)$ obtained by integrating $\Delta\phi$ between 0 and π, presented in the previous section, are fitted to the cluster model to allow an easier comparison to other measurements that provide cluster fit parameters.

Following the same approach as the PHOBOS and CMS Collaborations, described in section 2.2.1, the pseudorapidity correlation function is inter-

preted in the context of an independent cluster emission model using equation (2.7).

In order to extract the cluster properties from each of the data and MC $R(\Delta\eta)$ distributions presented in section 9.2, the following steps are carried out:

- First, the background distribution B as a function of $\Delta\eta$ (integrated over the full $\Delta\phi$ range) is corrected following the iteration procedure described in section 7.1.2.

- Second, the slope of $B(\Delta\eta)$ is parameterised with a straight line fit: $p[0]_B + p[1]_B\,\Delta\eta$.

- Third, the pseudorapidity correlation function $R(\Delta\eta)$ is fitted with equation (2.7), where $B(\Delta\eta)$ is given by the polynomial obtained in the second step and $\Gamma(\Delta\eta)$ is shown in equation (2.8). The exact functional form used to fit the curves is

$$p[0]\left[\frac{p[1]\ \exp\left(\frac{-(\Delta\eta-p[2])^2}{2\,p[3]^2}\right)}{p[0]_B + p[1]_B\,\Delta\eta} - 1\right], \tag{9.1}$$

  where $p[0] = \alpha$, $p[1]$ is just a proportionality constant, $p[2]$ and $p[3]$ are the mean and standard deviation of the gaussian, respectively, and $p[0]_B$ and $p[1]_B$ are the fixed values of the background linear fit.

- Finally, from the parameters in equation (9.1), the effective cluster size K_{eff} (equation (2.9)) can be calculated as

$$K_{\text{eff}} = p[0] + 1, \tag{9.2}$$

and the decay width as

$$\delta = \frac{p[3]}{\sqrt{2}}. \tag{9.3}$$

Figure 9.18 shows the corrected background distribution $B(\Delta\eta)$ at 7 TeV for events with $n_{ch} \geq 2$. The behaviour of the distribution can be parameterised as a straight line in $\Delta\eta$. This is also the case for distributions at 900 GeV and with $n_{ch} \geq 20$.

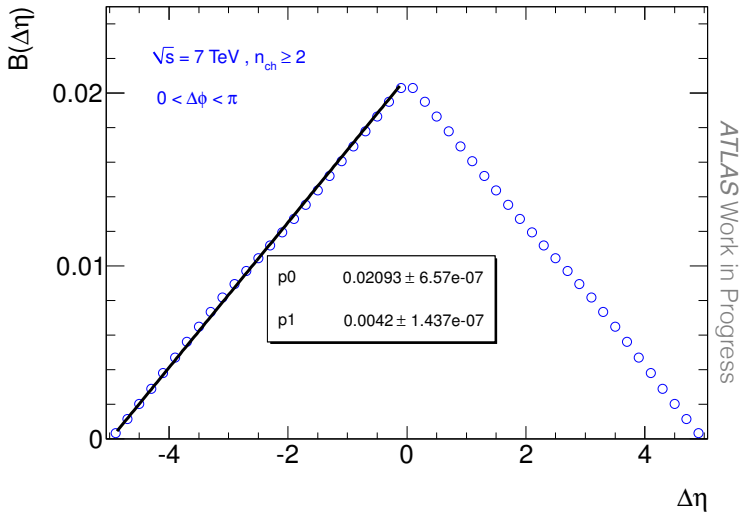


Figure 9.18: Corrected background distribution in $\Delta\eta$, obtained by integrating $\Delta\phi$ between 0 and π, for events with $n_{ch} \geq 2$ at 7 TeV. The behaviour of the background can be described by the parameters of a linear equation.

The data $R(\Delta\eta)$ distributions fitted with equation (9.1), similar to those from CMS in figure 2.9, are shown in figure 9.19. The distributions between the two experiments are in agreement with each other. As can be seen from these figures, the cluster model does not provide a good fit to the data. The most central $\Delta\eta$ bins are always above the fit curve, the width of the central peak is not correctly reproduced and the tail region is predicted to be flat. Even though the values for $\chi^2/\mathrm{N_{ndof}}$ of the fits are not (so) bad, it is difficult to believe that something can be concluded from these fits. However, previous experiments have argued that equation (9.1) is just a simple way to quanti-

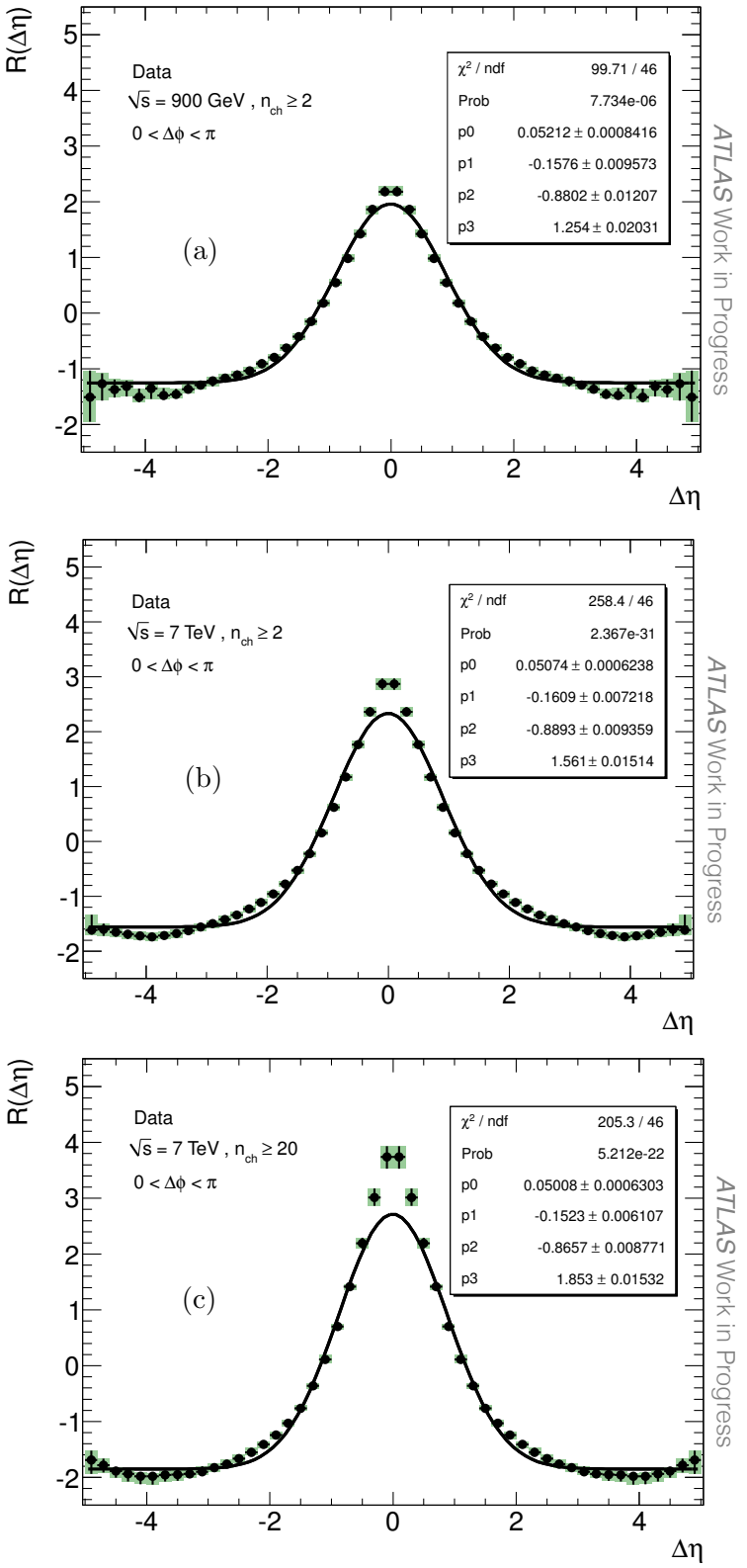


Figure 9.19: Cluster fits to the corrected data pseudorapidity correlation functions at (a) 900 GeV and 7 TeV for events with (b) $n_{ch} \geq 2$ and (c) $n_{ch} \geq 20$.

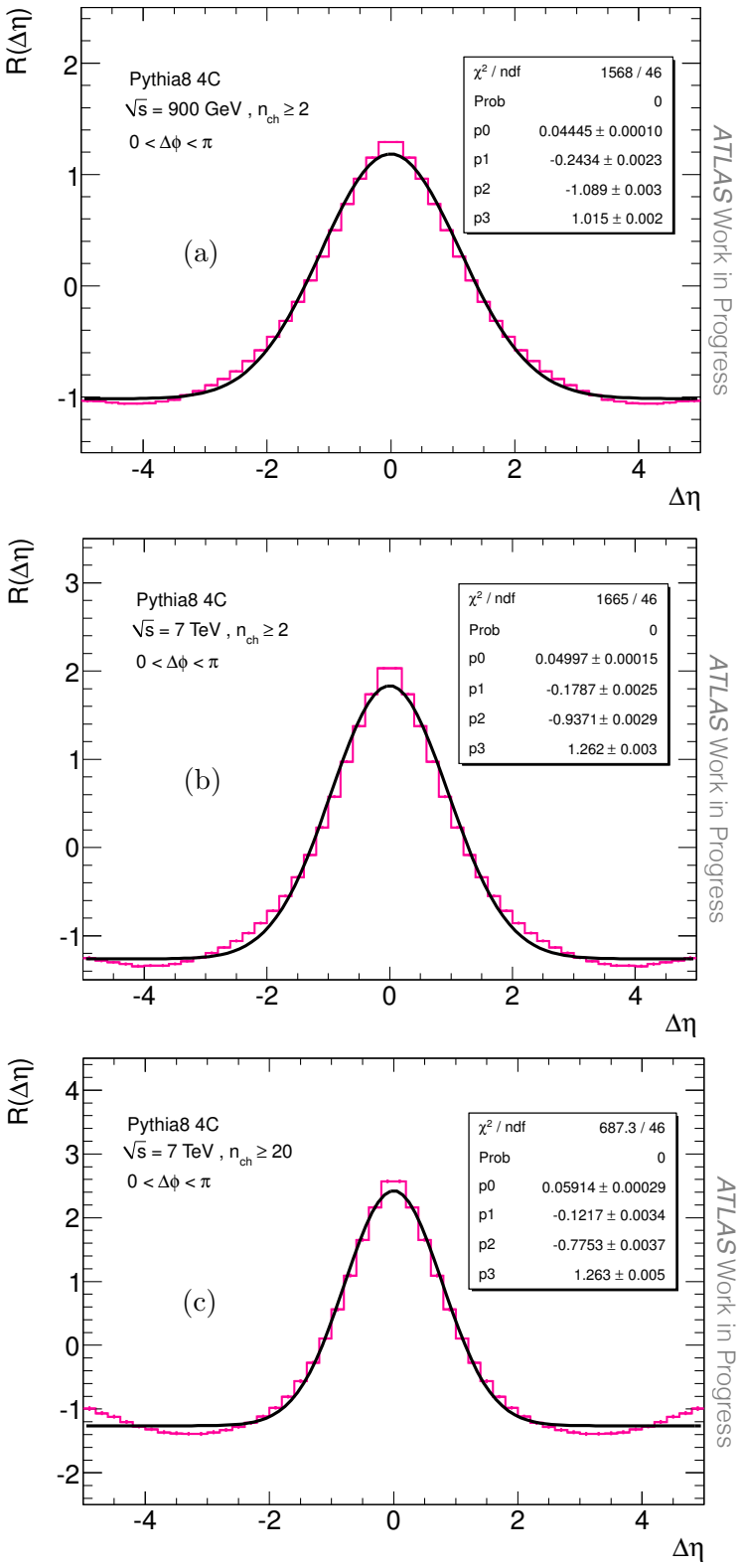


Figure 9.20: Cluster fits to the pseudorapidity correlation functions of tune 4C of PYTHIA8 at (a) 900 GeV and 7 TeV for events with (b) $n_{ch} \geq 2$ and (c) $n_{ch} \geq 20$.

tatively parameterise the correlation function as opposed to a precise test of a particular model [99]. For the sake of argument, the cluster properties are extracted from these fits and are summarised in table 9.1.

The cluster model does not fit any of the MC distributions either and in fact, the $\chi^2/\mathrm{N_{ndof}}$ of the fits is much worse. Figure 9.20 shows as an example the fitted distributions for the tune 4C of PYTHIA8 at both energies and particle multiplicities. The complete list of cluster properties from all MC samples is given in table 9.2.

For an easier visualisation of the results, in figures 9.21 and 9.22 the effective cluster size and the decay width, respectively, are plotted as a function of the centre-of-mass energy for data and MC. The solid markers correspond to measurements from the $n_{ch} \geq 2$ phase-space, while the higher-multiplicity results at 7 TeV are shown as open markers. In data (circles), the effective cluster size increases with both energy and multiplicity, while the decay width stays constant with energy and decreases slightly for events with $n_{ch} \geq 20$. The same trends for the cluster size are also seen in all the MC tunes, with the exception of HERWIG++ which predicts a smaller cluster size for high-multiplicities. Also in MC, the decay width decreases more pronouncedly. The conclusion reached from the data-MC comparisons in the previous section still holds: the MC tunes do not reproduce the values observed in data and HERWIG++ is the most discrepant tune.

Phase-Space	K_{eff}	δ	$\chi^2/\mathrm{N_{ndof}}$
900 GeV	2.2540 ± 0.0203	0.6224 ± 0.0086	2.1676
7 TeV ; $n_{ch} \geq 2$	2.5610 ± 0.0151	0.6288 ± 0.0066	5.6174
7 TeV ; $n_{ch} \geq 20$	2.8530 ± 0.0153	0.6121 ± 0.0062	4.4630

Table 9.1: Cluster properties for data in all phase-space regions.

MC Tune	Phase-Space	K_{eff}	Decay width δ	$\chi^2/\text{N}_{\text{ndof}}$
AMBT2B	900 GeV	1.8703 ± 0.0013	0.6885 ± 0.0011	65.9348
	7 TeV ($n_{ch} \geq 2$)	2.1430 ± 0.0010	0.6193 ± 0.0008	117.8479
	7 TeV ($n_{ch} \geq 20$)	2.6330 ± 0.0030	0.5818 ± 0.0011	72.6522
4C	900 GeV	2.0150 ± 0.0020	0.7700 ± 0.0021	34.0869
	7 TeV ($n_{ch} \geq 2$)	2.2620 ± 0.0030	0.6626 ± 0.0021	36.1957
	7 TeV ($n_{ch} \geq 20$)	2.2630 ± 0.0050	0.5482 ± 0.0026	14.9413
DW	900 GeV	1.6715 ± 0.0012	0.6718 ± 0.0013	25.6304
	7 TeV ($n_{ch} \geq 2$)	1.6746 ± 0.1467	0.5076 ± 0.0762	0.0146
	7 TeV ($n_{ch} \geq 20$)	1.8668 ± 0.1589	0.4371 ± 0.0535	0.0733
Perugia 2011	900 GeV	1.9723 ± 0.0015	0.7347 ± 0.0014	42.8696
	7 TeV ($n_{ch} \geq 2$)	2.1510 ± 0.0020	0.6265 ± 0.0014	46.1304
	7 TeV ($n_{ch} \geq 20$)	2.5150 ± 0.0040	0.5718 ± 0.0017	28.2826
HERWIG++	900 GeV	4.1960 ± 0.0030	0.8952 ± 0.0007	410.2174
	7 TeV ($n_{ch} \geq 2$)	4.5140 ± 0.0040	0.8202 ± 0.0011	518.6957
	7 TeV ($n_{ch} \geq 20$)	4.3240 ± 0.0070	0.8096 ± 0.0021	199.6304

Table 9.2: Cluster properties for the different MC tunes in all phase-space regions.

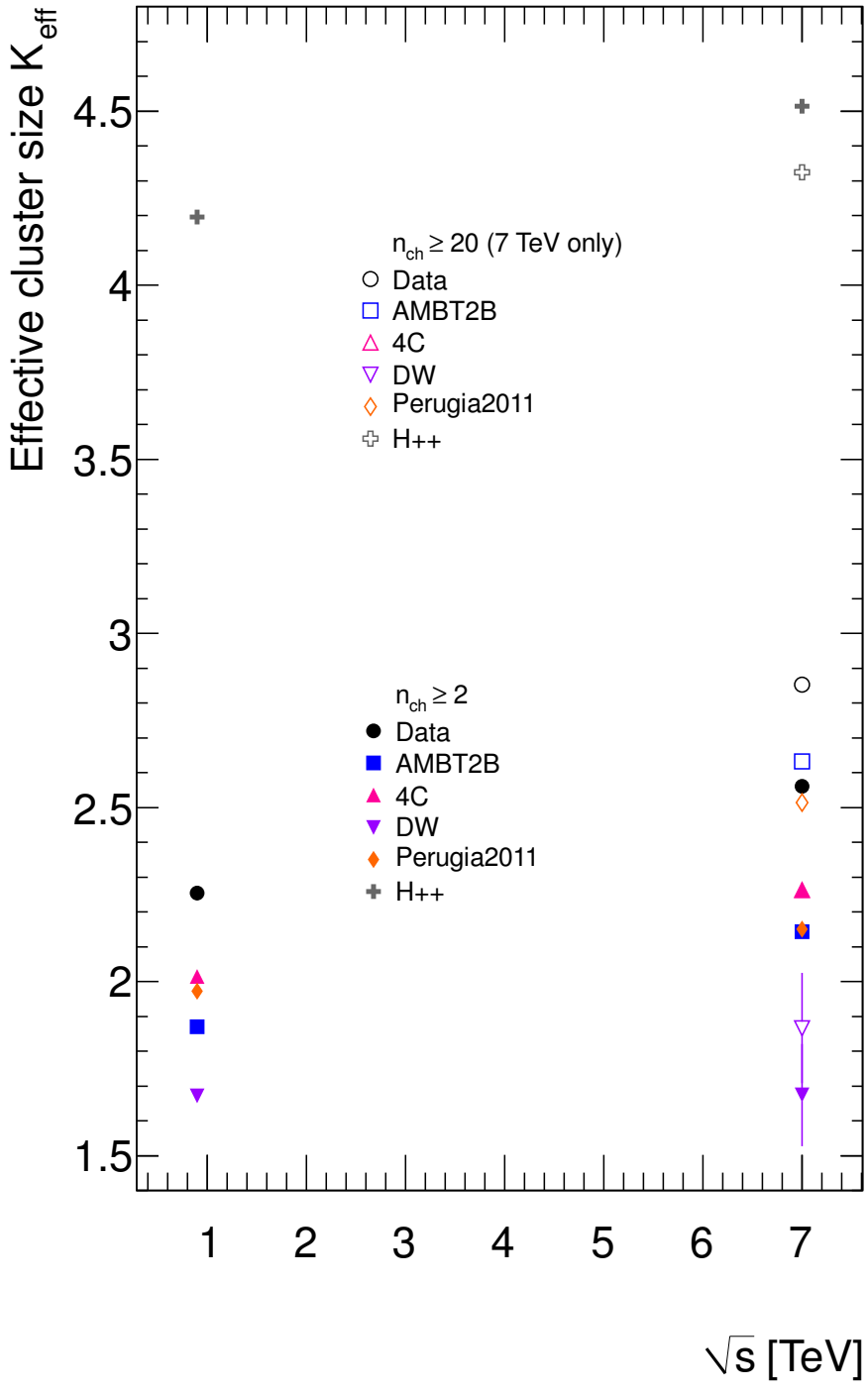


Figure 9.21: Effective cluster size for data and the different MC tunes as a function of the centre-of-mass energy. Solid (open) markers correspond to the $n_{ch} \geq 2$ ($n_{ch} \geq 20$)) phase-space.

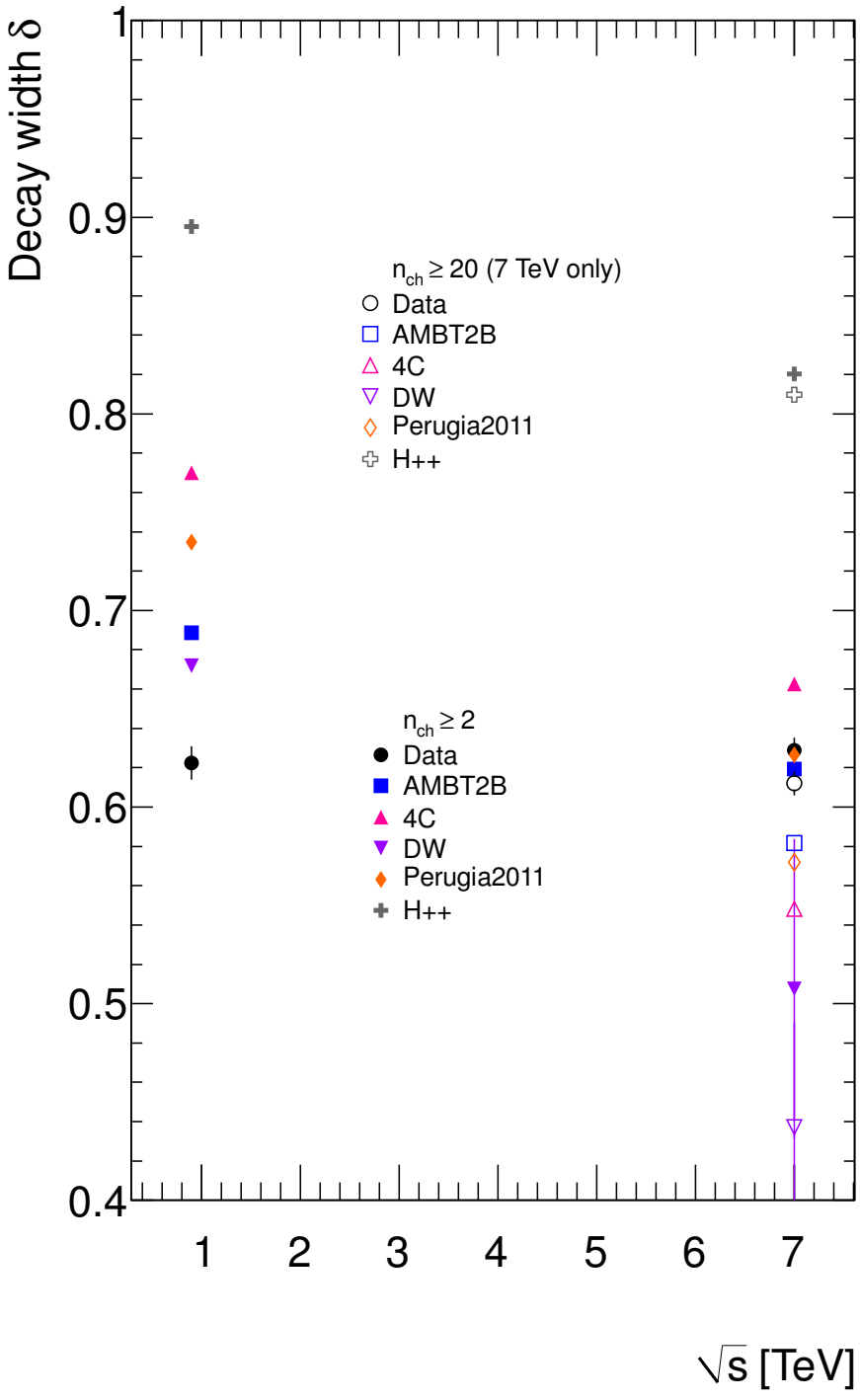


Figure 9.22: Cluster decay width for data and the different MC tunes as a function of the centre-of-mass energy. Solid (open) markers correspond to the $n_{ch} \geq 2$ ($n_{ch} \geq 20$)) phase-space.

Chapter 10

Conclusions

In little over two years of operation, the ATLAS experiment at the LHC has collected around 5.3 fb^{-1} of proton-proton collision data. From this sample, data collected during the early LHC runs is ideal to perform minimum bias measurements due to the low luminosity and, consequently, low number of multiple interactions per bunch crossing.

The inclusive two-particle angular correlation function was measured on a sample of minimum bias events collected with the ATLAS detector at $\sqrt{s} =$ 900 GeV and 7 TeV. The approximate integrated luminosities analysed were 7 μb^{-1} and 190 μb^{-1} for 900 GeV and 7 TeV, respectively. Correlations were measured for charged particles in the kinematic range $p_{\text{T}} > 100$ MeV and $|\eta| < 2.5$ in events with a particle multiplicity $n_{ch} \geq 2$. At 7 TeV, a second particle multiplicity event selection requirement was explored, namely $n_{ch} \geq 20$ in which the contribution from diffractive events is suppressed. These phase-space regions match the ones studied by the ATLAS Minimum Bias analysis [84] and as such, provide a nice complement for tuning purposes.

In order to unfold the detector effects from the true physics in this measurement, a novel model-independent correction method, the *Hit Backspace Once More* (HBOM) Method [92], was developed. This data-driven proce-

dure[1] was tested in Monte Carlo and good agreement was found between the generated and corrected distributions. The remaining differences were found to be associated to the presence of secondary particles. This, however, does not constitute a limitation of the correction method but rather of the detector model. The non-closure and the track reconstruction efficiency were found to be the dominant sources of systematic uncertainty.

The corrected two-particle correlation function exhibited a complex structure in $\Delta\eta$ and $\Delta\phi$ at both energies. This structure was dominated by a sharp peak around (0,0), attributed to particles originating from the same high-p_{T} process. This peak became more pronounced as both the energy and the particle multiplicity of the events increased, indicating the emergence of jet-like structures within the minimum bias event selection.

To explore in more detail the contributions from different underlying processes, the two-dimensional distributions were projected along both $\Delta\eta$ and $\Delta\phi$. These one-dimensional distributions allow for an easier comparison to the different Monte Carlo models and tunes. None of the MC models reproduced the strength of the correlations seen in the data and only in some of the projections they were able to correctly model the shape of the distributions. For the $n_{ch} \geq 2$ phase-space, the tune that performed better overall was 4C of PYTHIA8, while for the higher particle multiplicity events, the tunes of PYTHIA6, AMBT2B and Perugia 2011, were the closest to the data. In all distributions, HERWIG++ was the most discrepant tune. One possible explanation for the behaviour of HERWIG++ is its lack of diffractive components. However, the $n_{ch} \geq 20$ distributions contain little or no diffractive contribution

[1]The method is not entirely data-driven since the track reconstruction efficiency used to throw away tracks was determined from Monte Carlo. But this is the only place where simulation enters the correction method.

and HERWIG++ is still not able to describe them, which could mean that the other differences between PYTHIA and HERWIG++ (e.g. hadronisation model, approach to multiple parton interactions) could also be a contributing factor to the observed discrepancies.

Finally, the strength of the pseudorapidity correlation function $R(\Delta\eta)$, obtained by integrating $\Delta\phi$ between 0 and π, was interpreted in terms of a cluster emission model. It was found that this model does not provide a good fit to the data, neither around the central peak nor in the tail regions. To provide numbers that can be compared to measurements from previous experiments, the cluster fits were done to the data to extract cluster parameters. It was found that the effective cluster size increases with centre-of-mass energy and that the decay width is approximately constant, decreasing only slightly with energy. A direct comparison of the calculated cluster properties to the results from previous experiments cannot be done since the results presented in this thesis would need to be extrapolated to the same kinematic region. However, a qualitative comparison shows that, as illustrated in figure 2.8 in chapter 2, the same trends have been observed by previous experiments. In fact, in the case of the CMS experiment, with which the kinematic cuts are not that different, the values for cluster size and decay width are consistent.

Angular correlations between charged particles provide useful information and are sensitive to a wide range of parameters in the Monte Carlo simulations. They can give useful input for tuning purposes and for characterising the dynamics of soft interactions that could help discriminate between the available models. These results show that, despite many successes and qualitative agreement with experimental data, the phenomenology of soft interactions still requires some improvement, which may well extend beyond simply re-tuning experimental models.

Appendices

Appendix A

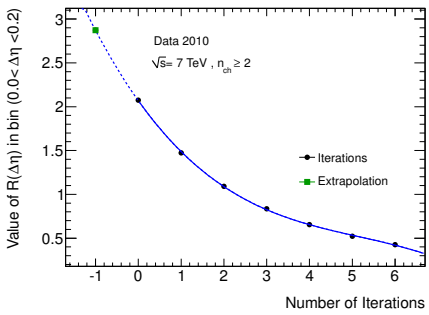
Third-degree Polynomial Fits

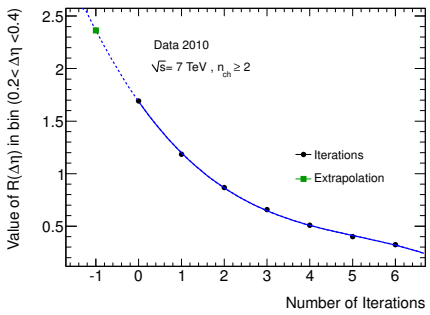
The complete set of polynomial fits to obtain the corrected pseudorapidity correlation distribution[1] for 7 TeV data, for events with $n_{ch} \geq 2$, are shown in figure A.1. The distribution itself has 50 bins but given that it is symmetric around $\Delta\eta = 0$, only 25 bins are given.

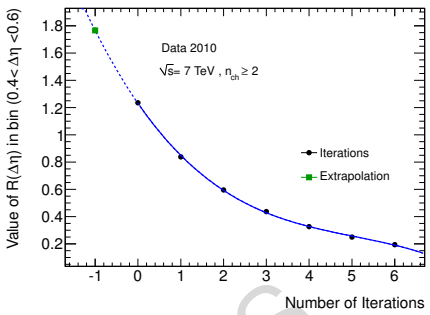
In addition, for illustration purposes, 10 bins of the two-dimensional correlation function $R(\Delta\eta, \Delta\phi)$, which are mainly bins with low-statistics, are shown in figure A.2 for the same data sample.

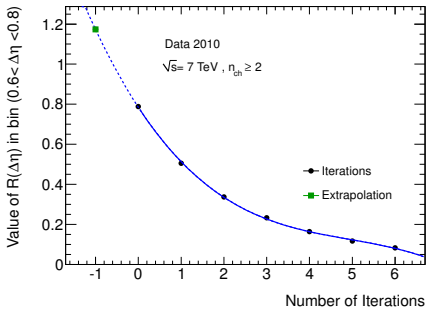
In all cases, a third degree polynomial is a good description of the behaviour of the observable in the track removal correction method.

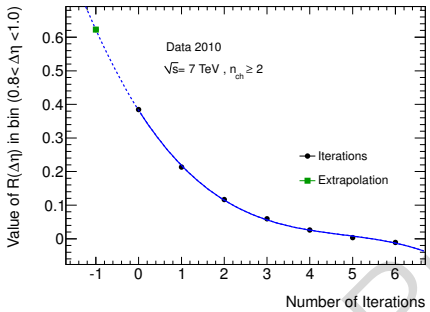
[1] Obtained by integrating $\Delta\phi$ between 0 and π.

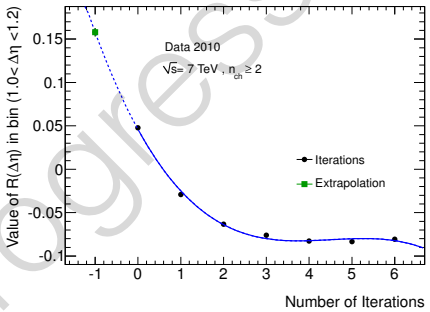


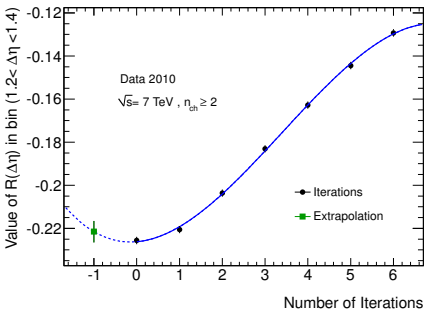


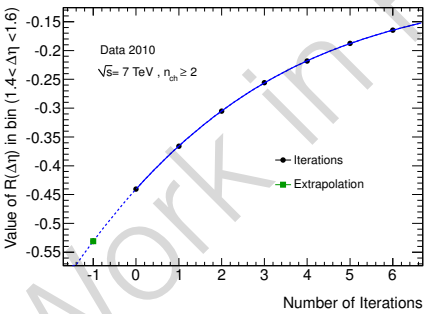


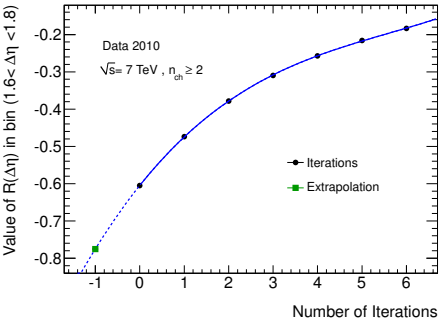


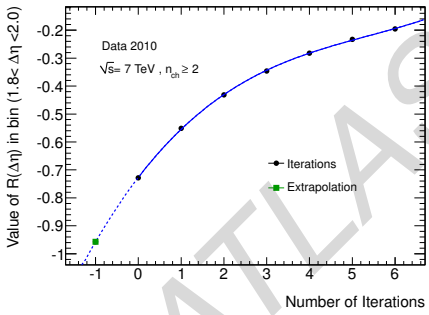


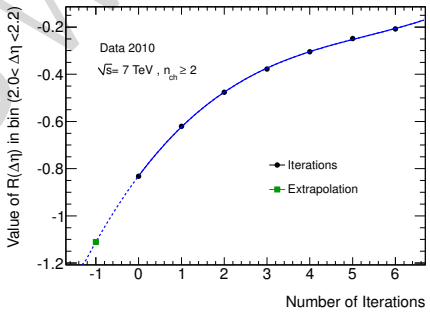


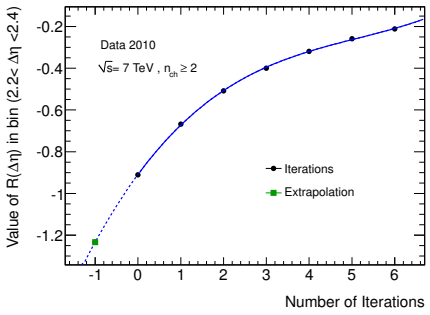


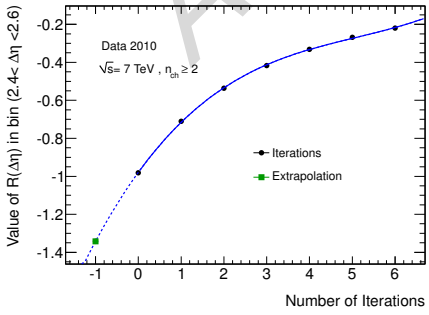


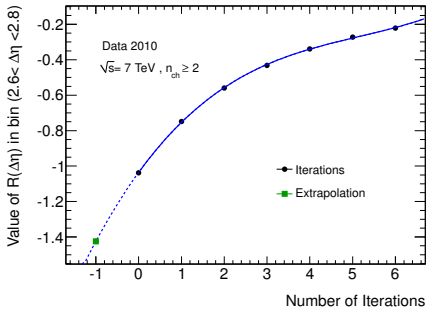


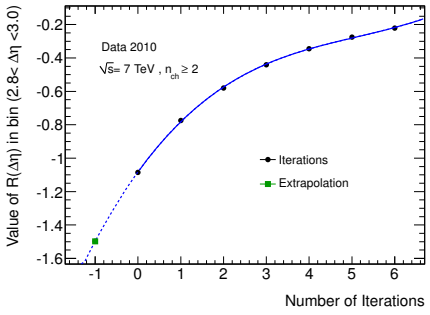












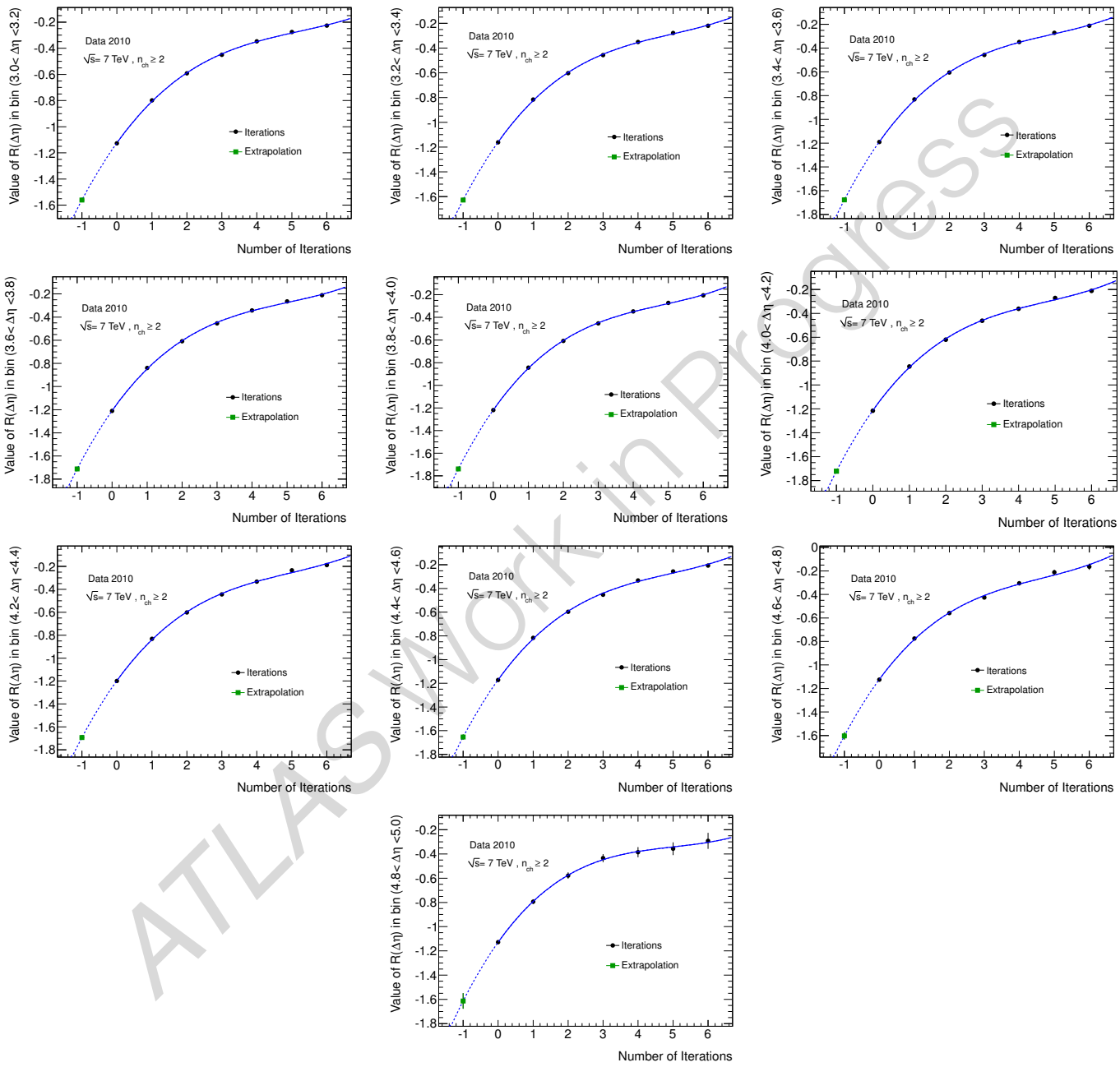


Figure A.1: Third-degree polynomial fits to all bins of the $R(\Delta\eta)$ distribution for data at 7 TeV ($n_{ch} \geq 2$). In total, $R(\Delta\eta)$ has 50 bins but, since it is symmetric by construction around $\Delta\eta = 0$, only 25 bins are shown.

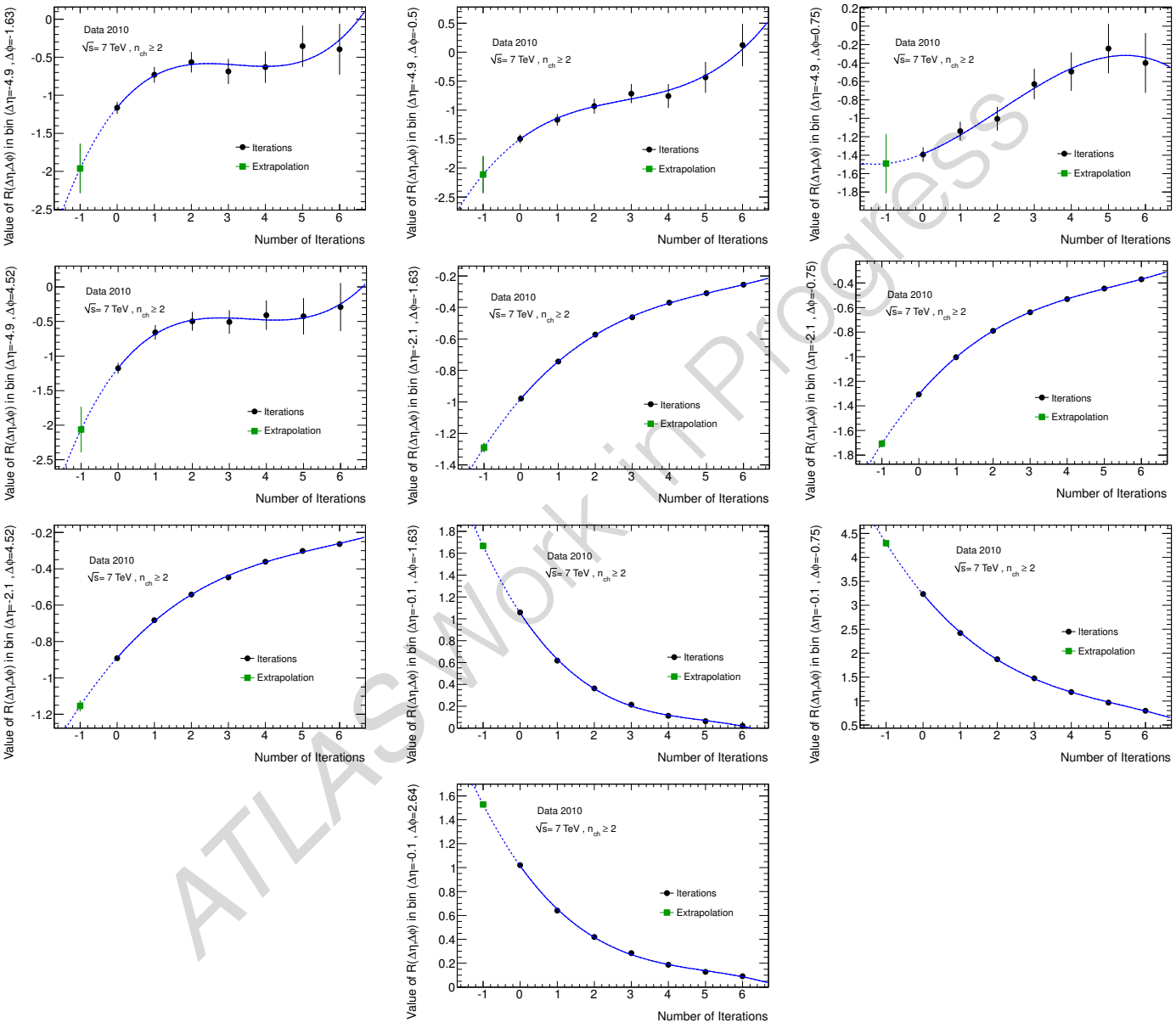


Figure A.2: Example of third-degree polynomial fits done to 10 bins of the $R(\Delta\eta, \Delta\phi)$ distribution for data at 7 TeV ($n_{ch} \geq 2$).

Appendix B

Non-Closure Uncertainties for 1D Projections

The size of the non-closure systematic uncertainty depends on the projection under study. Each distribution exhibits different behaviours; in some cases a flat uncertainty provides a reasonable description, while in some others linear parameterisations are needed. To illustrate this, the complete set of distributions used to extract the uncertainties for each projection in the $n_{ch} \geq 2$ phase-space at 7 TeV are presented in figures B.1 to B.5. The calculated values are summarised in table B.1.

Projection		Non-Closure Systematic Uncertainty
$R(\Delta\eta)$	$0 < \Delta\phi < \pi$	flat 0.05 for all $R(\Delta\eta)$
	$0 < \Delta\phi < \pi/2$	flat 0.05 for all $R(\Delta\eta)$
	$\pi/2 < \Delta\phi < \pi$	$0.0029 + 0.0865\, R(\Delta\eta)$
$R(\Delta\phi)$	$0 < \Delta\eta < 2$	$0.0804 - 0.0680\, R(\Delta\phi)$
	$2 < \Delta\eta < 5$	$-0.0863 - 0.0510\, R(\Delta\phi)$

Table B.1: Non-closure systematic uncertainties for the different 1D projections at 7 TeV ($n_{ch} \geq 2$).

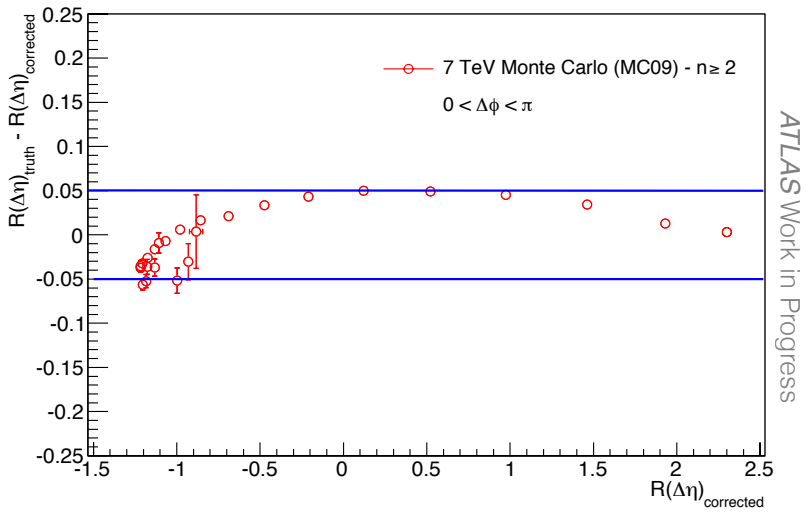


Figure B.1: Estimation of the non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between 0 and π for 7 TeV with $n_{ch} \geq 2$.

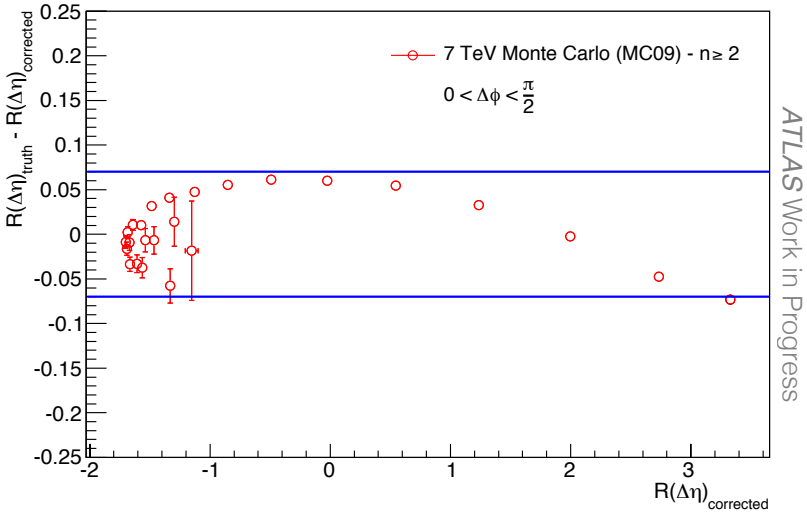


Figure B.2: Estimation of the non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between 0 and $\frac{\pi}{2}$ for 7 TeV with $n_{ch} \geq 2$.

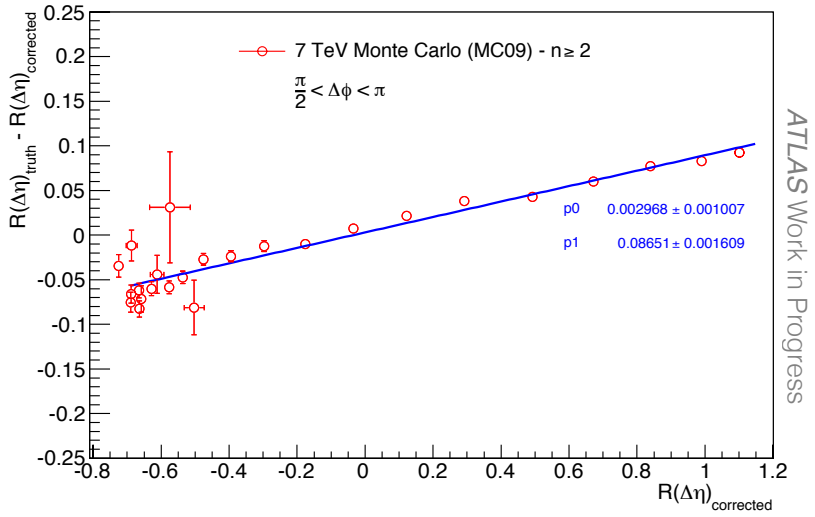


Figure B.3: Estimation of the non-closure systematic uncertainty for $R(\Delta\eta)$ where $\Delta\phi$ is integrated between $\frac{\pi}{2}$ and π for 7 TeV with $n_{ch} \geq 2$.

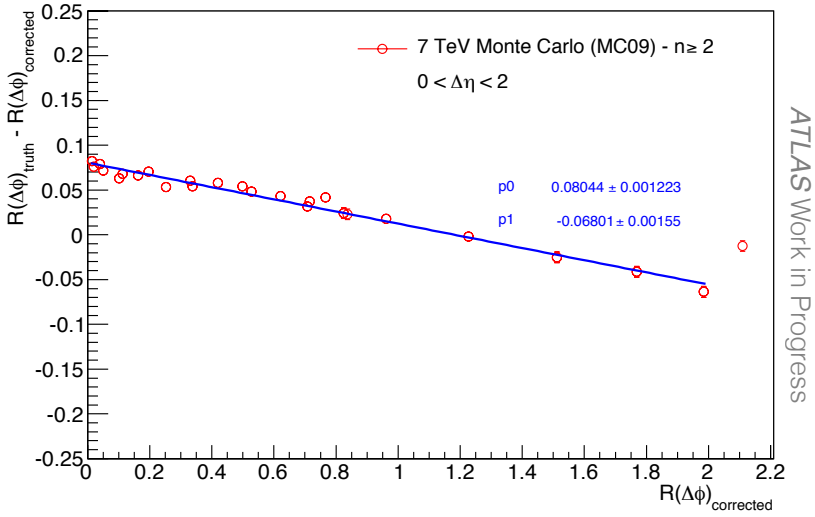


Figure B.4: Estimation of the non-closure systematic uncertainty for $R(\Delta\phi)$ where $\Delta\eta$ is integrated between 0 and 2 for 7 TeV with $n_{ch} \geq 2$.

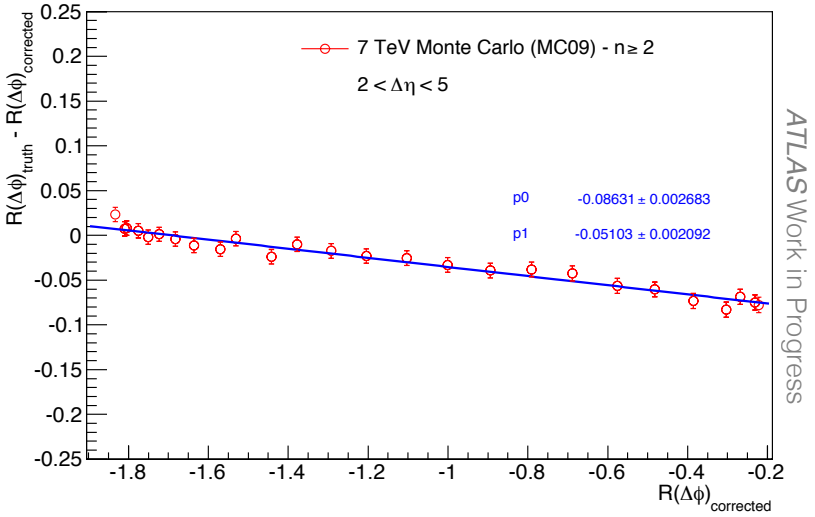


Figure B.5: Estimation of the non-closure systematic uncertainty for $R(\Delta\phi)$ where $\Delta\eta$ is integrated between 2 and 5 for 7 TeV with $n_{ch} \geq 2$.

Appendix C

Uncertainties for the 2D Distributions

C.1 Non-Closure Systematics

The estimation of the systematic uncertainties associated to the non-closure of the correction method for the two-dimensional distributions follows the same procedure applied, in section 8.2.1 and appendix B, to the $\Delta\eta$ and $\Delta\phi$ projections. The difference between the truth and corrected distributions is plotted as a function of $R(\Delta\eta, \Delta\phi)_{corr}$. As shown in figure C.1, for the measurement at 900 GeV a flat uncertainty of 0.2 is assigned to all bins. For the 7 TeV cases, for low values of $R(\Delta\eta, \Delta\phi)_{corr}$ the non-closure has an approximately flat behaviour. However, above a certain R value, the systematic uncertainty can be parameterised by means of a linear fit to the MC points. For the sample with $n_{ch} \geq 2$ events (figure C.2), a flat uncertainty of 0.1 is assigned to all values of $R(\Delta\eta, \Delta\phi)_{corr}$ smaller than 3; values greater than 3 are given by

$$0.2107 - 0.0925\, R(\Delta\eta, \Delta\phi).$$

Figure C.3 shows that for $n_{ch} \geq 20$, values of R smaller than 2 are better described by a flat uncertainty of 0.2, and everywhere else by

$$0.1712 - 0.0912\, R(\Delta\eta, \Delta\phi).$$

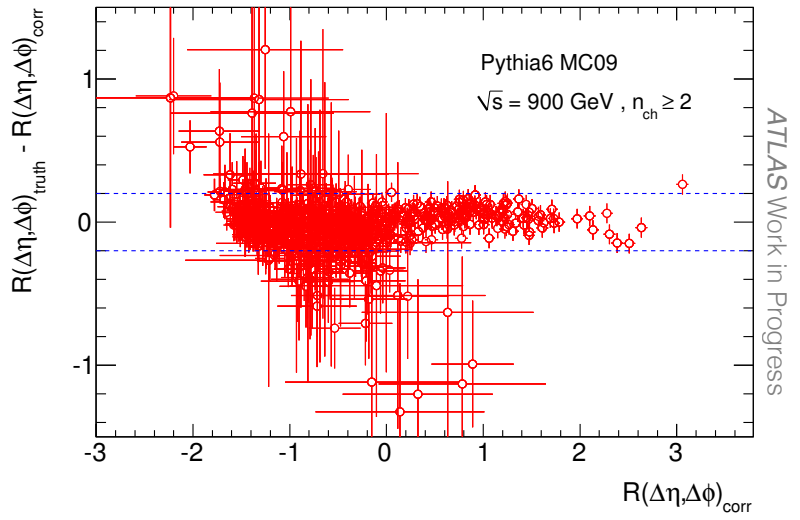


Figure C.1: Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 900 GeV Monte Carlo (MC09).

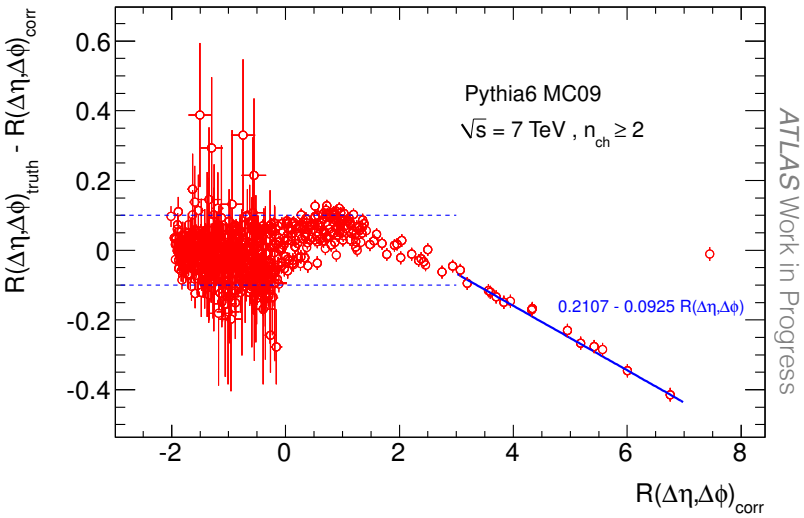


Figure C.2: Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 7 TeV Monte Carlo (MC09) for events with $n_{ch} \geq 2$.

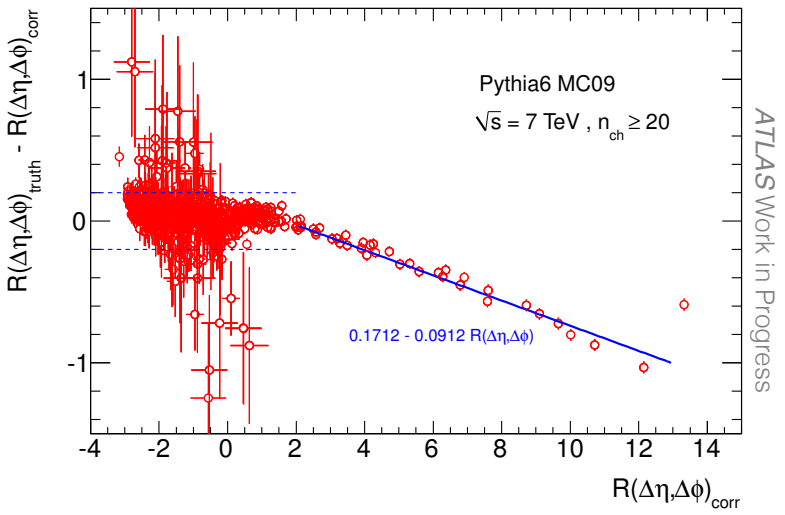


Figure C.3: Difference between the generated and corrected 2D correlation distributions as a function of $R(\Delta\eta, \Delta\phi)_{corr}$ in 7 TeV Monte Carlo (MC09) for events with $n_{ch} \geq 20$.

C.2 Total Uncertainties

Figures C.4 to C.6 show the total uncertainty in each bin of the two-dimensional correlation distributions for data at both energies and regions of phase-space. This total uncertainty is calculated by adding in quadrature the statistical and systematic uncertainties obtained following the methodology described in chapter 8. The largest contributions come from the non-closure of the correction method and the track reconstruction efficiency.

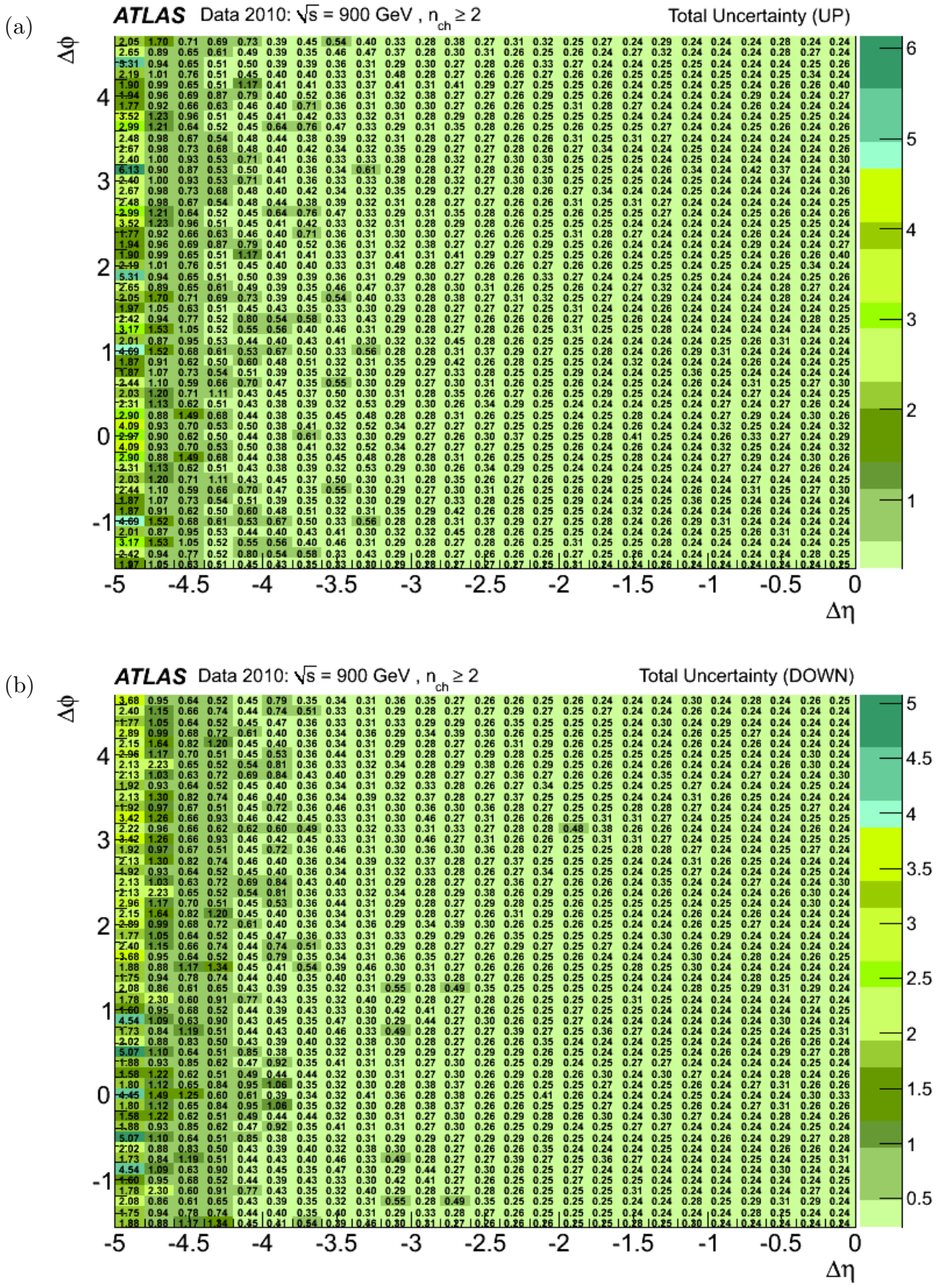


Figure C.4: Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 900 GeV. Since the correlation function is symmetric around $\Delta\eta = 0$ by construction, only the $\Delta\eta$ range between -5 and 0 is shown.

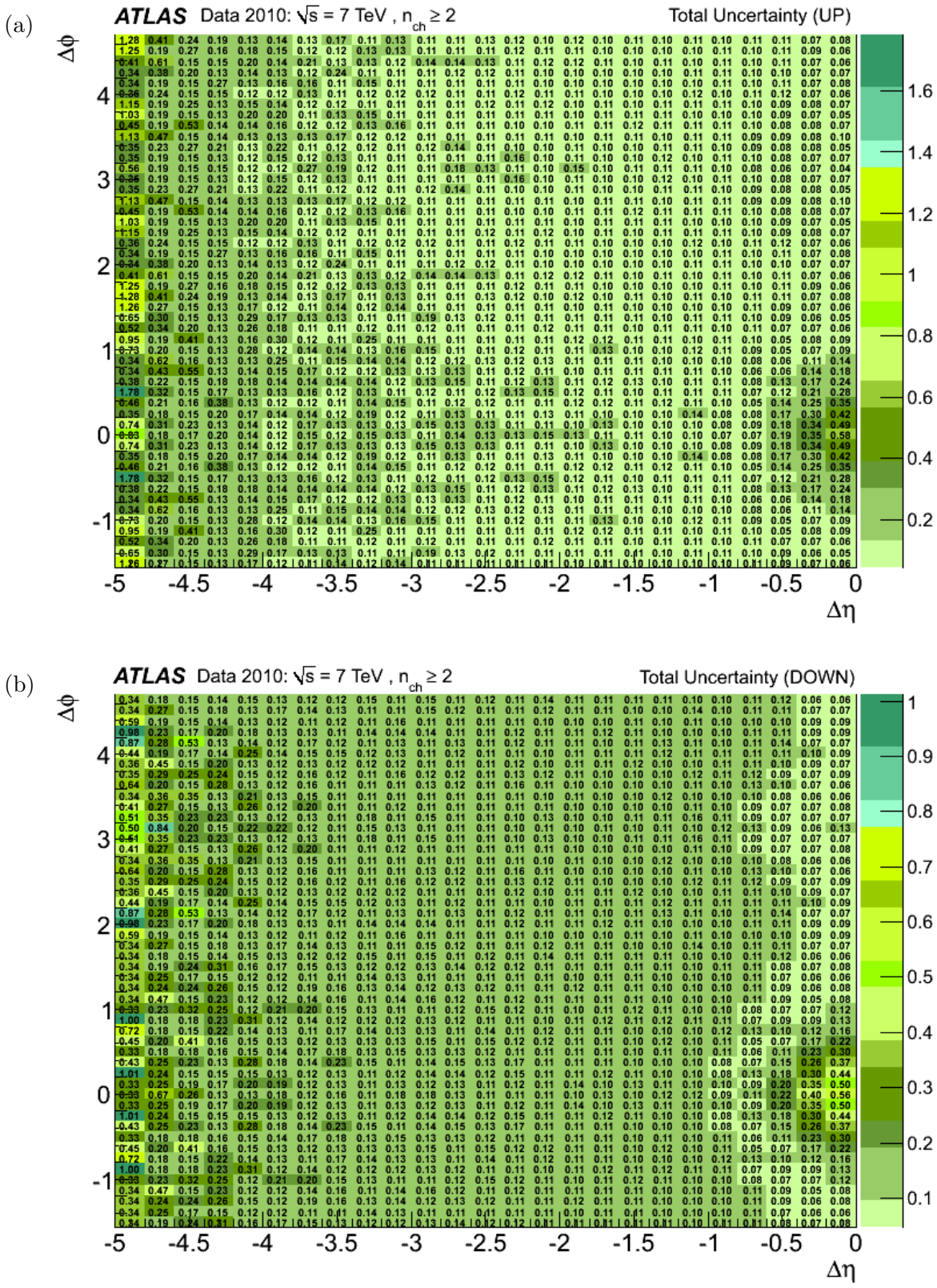
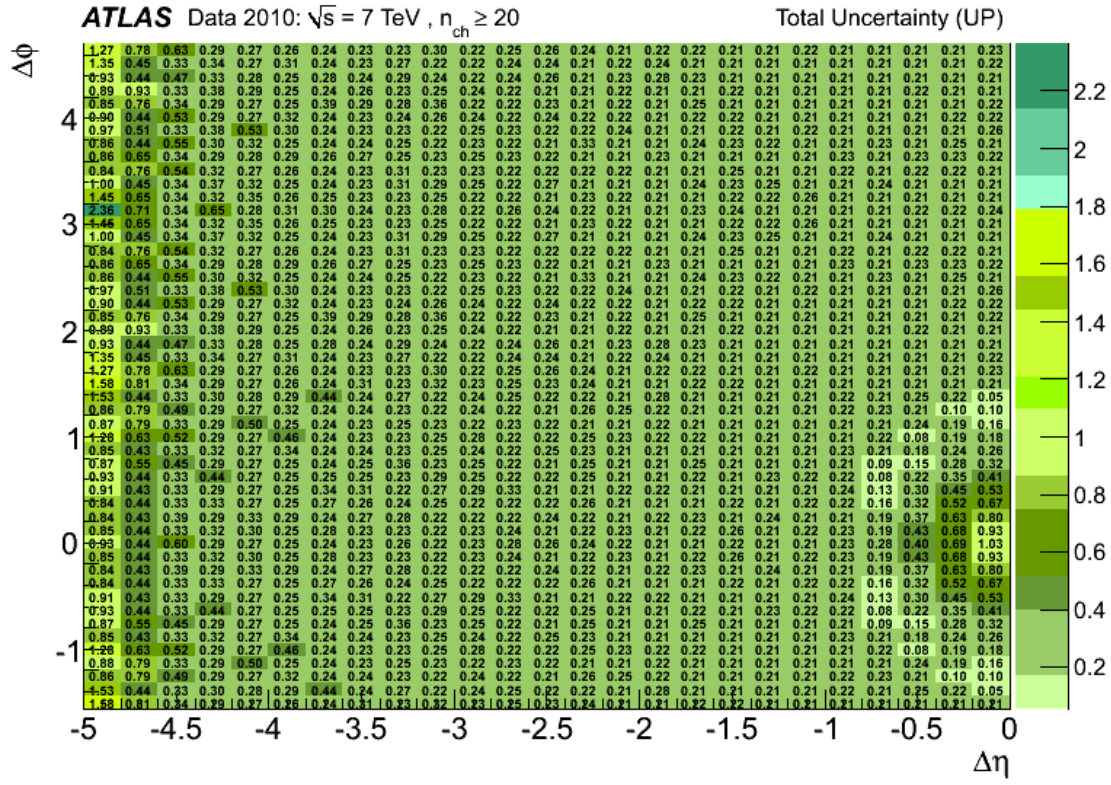


Figure C.5: Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 7 TeV for events with $n_{ch} \geq 2$. Since the correlation function is symmetric around $\Delta\eta = 0$ by construction, only the $\Delta\eta$ range between -5 and 0 is shown.

(a)



(b)

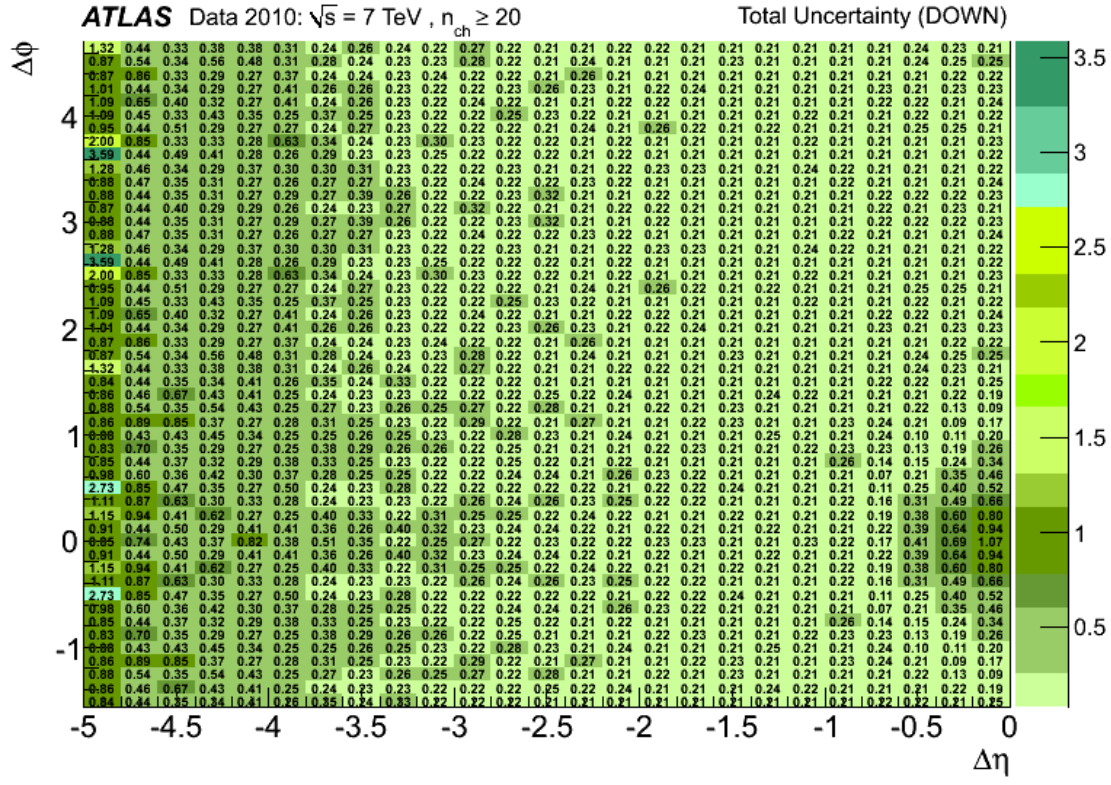


Figure C.6: Total uncertainty for each bin in $R(\Delta\eta, \Delta\phi)$ at 7 TeV for events with $n_{ch} \geq 20$. Since the correlation function is symmetric around $\Delta\eta = 0$ by construction, only the $\Delta\eta$ range between -5 and 0 is shown.

Bibliography

[1] D. H. Perkins, **Introduction to High Energy Physics**, Addison-Wesley Publishing Company (1987) 1–27.

[2] F. Halzen and A. Martin, **Quarks and Leptons: An Introductory Course in Modern Particle Physics**, John Wiley & Sons (1984) 1–32.

[3] Particle Data Group Collaboration, K. Nakamura et al., **Review of Particle Physics**, Journal of Physics G37 (2010) 075021.

[4] S. Glashow, **Partial Symmetries of Weak Interactions**, Nuclear Physics 22 (1961) 579–588.

[5] A. Salam, **Weak and electromagnetic interactions**, Elementary Particle Theory: Relativistic Groups and Analyticity, Almqvist & Wiksell, Stockholm (1968) 367–377.

[6] S. Weinberg, **A model of leptons**, Phys. Rev. Lett. 19 (1967) 1264.

[7] P. Higgs, **Broken Symmetries, Massless Particles and Gauge Fields**, Phys. Rev. Lett. 12 (1964) 132–133.

[8] P. Higgs, **Broken Symmetries and the Masses of the Gauge Bosons**, Phys. Rev. Lett. 13 (1964) 508–509.

[9] F. Englert and R. Brout, **Broken Symmetry and the Mass of the Gauge Vector Mesons**, Phys. Rev. Lett. 13 (1964) 321–323.

[10] The LEP Working Group for Higgs Boson Searches: ALEPH, DELPHI, L3 and OPAL Collaboration, **Search for the Standard Model Higgs Boson at LEP**, Phys. Lett. B565 (2003) 61–75.

[11] Tevatron NewPhenomena/Higgs Working Group: CDF and D0 Collaboration, **Combined CDF and D0 Upper Limits on Standard Model Higgs Boson Production with up to 8.6 fb^{-1} of Data**, `arXiv:1107.5518 [hep-ex]`.

[12] ATLAS Collaboration, **Combined Standard Model Higgs boson searches with up to 2.3 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV at the LHC**, ATLAS-CONF-2011-157**,** `http://cdsweb.cern.ch/record/1399599`.

[13] M. Gell-Mann, **A Schematic Model for Baryons and Mesons**, Phys. Lett. 8 (1964) 214–215.

[14] G. Zweig, **An SU(3) model for Strong Interaction Symmetry and its Breaking**, CERN Pre-Prints: 8182/TH.401, 8419/TH.412**,** `http://cdsweb.cern.ch/record/570209`.

[15] R. Ellis, W. Stirling, and B. Webber, **QCD and Collider Physics**, Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press (2003) 1–21.

[16] W. Murray (ed.), **Proceedings of the School for Experimental High Energy Physics Students**, Rutherford Appleton Laboratory, Technichal Report RAL-TR-2010-010 (2010) 115–145, `http://www.stfc.ac.uk/PPD/resources/pdf/StandardModel09.pdf`.

[17] TOTEM Collaboration, **First measurement of the total proton-proton cross-section at the LHC energy of $\sqrt{s} = 7$ TeV**, EPL 96 (2011) 21002.

[18] ATLAS Collaboration, **ATLAS high-level trigger, data-acquisition and controls: Technical Design Report**, ATLAS-TDR-016/CERN-LHCC-2003-022, `http://cdsweb.cern.ch/record/616089`.

[19] T. Sjostrand, S. Mrenna, and P. Skands, **PYTHIA 6.4 Physics and Manual**, JHEP 05 (2006) 026.

[20] M. Bahr et al., **Herwig++ Physics and Manual**, Eur. Phys. J. C58 (2008) 639–707.

[21] R. Engel, **Photoproduction within the two component dual parton model: Amplitudes and cross-sections**, Z. Phys. C66 (1995) 203–214.

[22] A. Capella, **Dual parton model**, Surveys in High Energy Physics 16 (2001) 175–186.

[23] B. Andersson et al., **A model for low-p_{T} hadronic reactions with generalizations to hadron-nucleus and nucleus-nucleus collisions**, Nucl. Phys B281 (1987) 289.

[24] M. Bähr et al., **Soft Interactions in Herwig++**, `arXiv:0905.4671`.

[25] B. Webber, **A QCD Model for Jet Fragmentation including Soft Gluon Interference**, Nucl. Phys. B238 (1984) 492.

[26] S. Gieseke et al., **Herwig++ 2.5 Release Note**, `arXiv:1102.1672`.

[27] T. Sj ostrand, S. Mrenna, and P. Skands, **Brief Introduction to Pythia 8.1**, Comput. Phys. Comm. 178 (2008) 852–867.

[28] M. Albrow et al., **Tevatron-for-LHC Report of the QCD Working Group**, arXiv:hep-ph/0610012.

[29] H. Lai et al., **Global QCD analysis of parton structure of the nucleon: CTEQ5 parton distributions**, Eur. Phys. J. C12 (2000) 375–392.

[30] T. Sj ostrand and P. Skands, **Transverse-momentum-ordered showers and interleaved multiple interactions**, Eur. Phys. J. C39 (2005) 129.

[31] P. Skands, **Tuning Monte Carlo Generators: The Perugia Tunes**, arXiv:1005.3457.

[32] ATLAS Collaboration, **ATLAS Monte Carlo Tunes for MC09**, ATL-PHYS-PUB-2010-002, http://cdsweb.cern.ch/record/1247375.

[33] ATLAS Collaboration, G. Aad et al., **ATLAS Tunes of Pythia6 and Pythia8 for MC11**, ATL-PHYS-PUB-2011-009, http://cdsweb.cern.ch/record/1363300.

[34] 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 a new PYTHIA6 tune**, ATLAS-CONF-2010-031, http://cdsweb.cern.ch/record/1277665.

[35] P. Skands, **The Perugia Tunes**, arXiv:0905.3418 [hep-ph].

[36] A. Sherstnev and R. Thorne, **Parton distributions for LO generators**, Eur. Phys. J. C55 (2008) 553–575.

[37] J. Pumplin et al., **New generation of parton distributions with uncertainties from global QCD analysis**, JHEP 07 (2002) 012.

[38] R. Corke and T. Sjostrand, **Interleaved parton showers and tuning prospects**, JHEP 03 (2011) 032.

[39] `http://projects.hepforge.org/herwig/trac/wiki/MB_UE_tunes`.

[40] W. Kittel and E. De Wolf, **Soft Multihadron Dynamics**, World Scientific Publishing (2005) 393–428.

[41] K. Eggert et al., **Angular Correlations between the Charged-Particles produced in pp Collisions at ISR Energies**, Nuclear Physics B86 (1975) 201–215.

[42] UA5 Collaboration, R. E. Ansorge et al., **Charged particle correlations in $\bar{p}p$ collisions at c.m. energies of 200, 546 and 900 GeV**, Z. Phys. C37 (1988) 191–213.

[43] CMS Collaboration, V. Khachatryan et al., **Observation of Long-Range, Near-Side Angular Correlations in Proton-Proton Collisions at the LHC**, JHEP 09 (2010) 091.

[44] D. Velicanu, **Ridge correlation structure in high-multiplicity pp collisions with CMS**, J. Phys. G: Nucl. Part. Phys. 38 (2011) 124051.

[45] A. Morel and G. Plaut, **How do clusters look in semi-inclusive cross sections?**, Nuclear Physics B78 (1974) 541–551.

[46] PHOBOS Collaboration, B. Alver et al., **Cluster properties from two-particle angular correlations in $p+p$ collisions at $\sqrt{s} = 200$ and 410 GeV**, Physical Review C75 (2007) 054913.

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

[48] http://public.web.cern.ch/public/en/research/AccelComplex-en.html.

[49] http://atlas.web.cern.ch/Atlas/public/EVTDISPLAY/events.html.

[50] CMS Collaboration, S. Chatrchyan et al., **The CMS Experiment at the CERN LHC**, JINST 3 (2008) S08004.

[51] LHCb Collaboration, A. Agusto Alves Jr et al., **The LHCb Detector at LHC**, JINST 3 (2008) S08005.

[52] TOTEM Collaboration, G. Anelli et al., **The TOTEM Experiment at the CERN Large Hadron Collider**, JINST 3 (2008) S08007.

[53] LHCf Collaboration, O. Adriani et al., **The LHCf Experiment at the CERN Large Hadron Collider**, JINST 3 (2008) S08006.

[54] ALICE Collaboration, K. Aamodt et al., **The ALICE Experiment at the CERN LHC**, JINST 3 (2008) S08002.

[55] ATLAS Collaboration, G. Aad et al., **The ATLAS Experiment at the CERN Large Hadron Collider**, JINST 3 (2008) S08003.

[56] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsATLASDetector.

[57] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults.

[58] ATLAS Collaboration, G. Duckeck (ed.) et al., **ATLAS Computing Technical Design Report**, ATLAS TDR–017, CERN-LHCC-2005-022, http://cdsweb.cern.ch/record/837738.

[59] G. Barrand et al., **Gaudi - A Software Architecture and Framework for building HEP Data Processing Applications**, Proceedings for the International Conference on Computing in High Energy Physics (CHEP) (2000) .

[60] J. Boudreau et al., **The GeoModel toolkit for detector description**, Proceedings for the International Conference on Computing in High Energy Physics (CHEP) (2004) .

[61] S. Spagnolo et al., **The Description of the ATLAS Detector**, Proceedings for the International Conference on Computing in High Energy Physics (CHEP) (2004) .

[62] ATLAS Collaboration, G. Aad et al., **The ATLAS Simulation Infrastructure**, Eur. Phys. J. C70 (2010) 823–874.

[63] M. Dobbs and J. Hansen, **The HepMC C++ Monte Carlo event record for high energy physics**, Comput. Phys. Commun. 134 (2001) 41–46.

[64] G. Agostellini et al., **Geant4 - a simulation toolkit**, Nucl. Instrum. Methods Phys. Res. A506 (2003) 250–303.

[65] J. Allison et al., **Geant4 developments and applications**, IEEE Trans. Nucl. Sci. 53 (2006) 270–278.

[66] E. Barberio et al., **The Geant4-based ATLAS fast electromagnetic shower simulation**, ATL-SOFT-CONF-2007-002, `http://cdsweb.cern.ch/record/1064665`.

[67] E. Barberio et al., **The fast shower simulation in the ATLAS calorimeter**, J. Phys. Conf. Ser. 119 (2007) 001001.

[68] E. Richter-Was, D. Froidevaux, and L. Poggioli, **ATLFAST 2.0 a fast simulation package for ATLAS**, ATL-PHYS-98-131, http://cdsweb.cern.ch/record/683751.

[69] http://www.hep.ucl.ac.uk/atlas/atlfast/.

[70] K. Edmonds, **The Fast ATLAS track simulation (FATRAS)**, ATL-SOFT-PUB-2008-01, http://cdsweb.cern.ch/record/1091969.

[71] F. Akesson et al., **The ATLAS Tracking Event Data Model**, ATL-SOFT-PUB-2006-004, http://cdsweb.cern.ch/record/973401.

[72] A. Salzburger et al., **The ATLAS Tracking Geometry Description**, ATL-SOFT-PUB-2007-004, http://cdsweb.cern.ch/record/1038098.

[73] A. Salzburger (ed.) et al., **The new ATLAS Track Reconstruction (NEWT)**, J. Phys.: Conf. Ser. 119 (2008) 032014.

[74] A. Salzburger et al., **Single Track Performance of the Inner Detector New Track Reconstruction (NEWT)**, ATL-INDET-PUB-2008-002, http://cdsweb.cern.ch/record/1092934.

[75] R. Frühwirth, **Application of Kalman Filtering to Track and Vertex Fitting**, Nucl. Inst. Meth. A262 (1987) 444–450.

[76] T. Cornelissen et al., **The global χ^2 track fitter in ATLAS**, J. Phys.: Conf. Ser. 119 (2008) 032013.

[77] R. Duda and P. Hart, **Use of the Hough Transformation to Detect Lines and Curves in Pictures**, Comm. ACM 15 (1972) 11–15.

[78] G. Piacquadio, K. Prokofiev, and A. Wildauer, **Primary Vertex Reconstruction in the ATLAS Experiment at the LHC**, J. Phys.: Conf. Ser. 119 (2008) 032033.

[79] A. Peisert and F. Sauli (ed.), **Instrumentation in High Energy Physics - Silicon microstrip detectors**, Advanced Series on Directions in High Energy Physics - Vol. 9, World Scientific Publishing (1992) 1–79.

[80] A. Ahmad et al., **The silicon microstrip sensors of the ATLAS semiconductor tracker**, Nucl. Inst. Methods Phys. Res. A578 (2007) 98–118.

[81] I. Gorelov et al., **A measurement of Lorentz angle and spatial resolution of radiation hard silicon pixel sensors**, Nucl. Instrum. Methods in Phys. Res. A481 (2002) 204–221.

[82] S. Haywood, **Local Coordinate Frames for the Alignment of Silicon Detectors**, ATL-INDET-2004-001, `http://cdsweb.cern.ch/record/705291`.

[83] CMS Collaboration, V. Khachatryan, **Charged particle multiplicities in pp interactions at $\sqrt{s}$ = 0.9, 2.36 and 7 TeV**, JHEP 01 (2011) 079.

[84] ATLAS Collaboration, G. Aad et al., **Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC**, New J. Phys. 13 (2011) 053033.

[85] R. Kwee, **Minimum Bias Triggers at ATLAS, LHC**, `arXiv:0812.0613`.

[86] M. Leyton, **Minimum Bias Measurements with the ATLAS Detector at the CERN Large Hadron Collider**, PhD Thesis, University of California (2009) 49–60.

[87] ATLAS Collaboration, **Performance of the Minimum Bias Trigger in p-p Collisions at $\sqrt{s}$ = 7 TeV**, ATLAS-CONF-2010-068, http://cdsweb.cern.ch/record/1281343.

[88] C. Ohm and T. Pauly, **The ATLAS beam pick-up based timing system**, Nucl. Instrum. Methods in Phys. Res. A623 (2010) 558–560.

[89] ATLAS Collaboration, **Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC**, ATLAS-CONF-2010-027, http://cdsweb.cern.ch/record/1277659.

[90] ATLAS Collaboration, **Charged-particle multiplicities in pp interactions for track $p_{\mathrm{T}} > 100$ MeV at $\sqrt{s} = 0.9$ and 7 TeV measured with the ATLAS detector at the LHC**, ATLAS-CONF-2010-046, http://cdsweb.cern.ch/record/1281296.

[91] ATLAS Collaboration, G. Aad et al., **Charged-particle multiplicities in pp interactions at $\sqrt{s}$ = 900 GeV measured with the ATLAS detector at the LHC**, Phys. Lett. B688 (2010) 21–42, arXiv:1003.3124.

[92] J. Monk and C. Oropeza-Barrera, **The HBOM Method for Unfolding Detector Effects**, arXiv:1111.4896.

[93] Private communications with James Monk.

[94] G. Cowan, **Statistical Data Analysis**, Oxford University Press (1998) 20–22.

[95] http://root.cern.ch/.

[96] P. L'Ecuyer, **Tables of Maximally-Equidistributed Combined LFSR Generators**, Mathematics of Computation 68 (1999) 225.

[97] M. Matsumoto and V. Nishimura, **Mersenne Twister: A 623-dimensional equidistributed uniform pseudorandom number generator**, ACM Transactions on Modelling and Computer Simulation 8 (1998) 3–30.

[98] ATLAS Collaboration, **Measurement of Inclusive Two-Particle Angular Correlations in Proton-Proton Collisions at $\sqrt{s}$ = 900 GeV and 7 TeV**, ATLAS-CONF-2011-055, http://cdsweb.cern.ch/record/1342570.

[99] CMS Collaboration, **Two-Particle Angular Correlations and Cluster Properties in pp Collisions at $\sqrt{s}$ = 0.9, 2.36 and 7 TeV**, CMS PAS QCD-10-002, http://cdsweb.cern.ch/record/1267376.