# Ph.D Thesis

November 5, 2010

# Charged particle multiplicities in pp interactions at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV measured with the ATLAS detector at the LHC.


Author: Matteo Volpi

*Institut de Física d'Altes Energies, Edifici Cn Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.*

Director: Ilya Korolkov Tutor: José Maria Crespo

# Contents

# Chapter 1

# Introduction

In the 20th century, particle physics experiments have proven crucial for our understanding of nature. Particle accelerators like the Large Hadron Collider (LHC) [1] boost subatomic particles to nearly the speed of light, before letting them collide.

The particles that are created in the collisions are detected by a particle detector. These detectors are extraordinarily complex, requiring years of research and development. The subject of this thesis is focused on how the ATLAS (A Toroidal LHC ApparatuS) detector is used to measure the central pseudorapidity and transverse momentum distributions of charged particles produced in inelastic proton-proton (pp) collisions during early running at the LHC.

The minimum bias events, to be defined in Section 1.1, allow the soft-part of the underlying event (UE) in high-$p_T$ collisions to be characterised. Studies of inclusive particle distributions in minimum bias events in pp collisions are important:

- to understand the physics behind the pp collisions, as well as to commission the ATLAS detector;
- to be able to unfold soft underline component present in any hard scattering final state. This holds for either the hypothetical final state of new physics or for perturbative quantum chromodynamics (QCD) [2], which constitutes major background to the searches of new physics;
- to provide the baseline for measurements in heavy-ion collisions, such as allowing differences in the number of particles to be attributed to QCD effects rather than the simple scaling of the number of nucleons.

The analysis presented here is done with very early data, from low-luminosity running of the LHC, when not more than one soft interaction per bunch collision was present. The studies reported here are compared on a full simulation of the ATLAS detector.

The outline of the thesis is as follows:

- Chapter 1 starts with a brief overview of the definitions in minimum bias measurement.
- Chapter 2 gives a overview of the LHC accelerator complex and the ATLAS detector. Focus is given to the components of the detector that are relevant for the analysis.

- Chapter 3 explains some of the theoretical models important for the measurement. Monte Carlo (MC) simulation of LHC data is presented and discussed.
- Chapter 4 describes how events for the analysis are triggered and selected. Efficiencies and acceptances of the trigger are evaluated using simulated data and real data.
- Chapter 5 describes how charged particle tracks are reconstructed and selected in ATLAS.
- Chapter 6 studies the trigger, vertex reconstruction and tracking efficiencies. Selection criteria for reconstructed tracks used in the analysis are also given.
- Chapter 7 gives an overview of the possible sources of background.
- Chapter 8 details the analysis procedure for applying corrections to the selected data.
- Chapter 9 addresses the sources of various systematic uncertainties on the measurement. An estimate is given for each source and a total systematic uncertainty is assigned.
- Chapter 10 shows the results and the comparison with the results from other experiments.

## 1.1 Definitions and corrections used in minimum bias measurements in ATLAS and in the previous studies

High-energy proton collisions at the LHC can be described in terms of parton interactions. However, our ability to describe parton scatterings through QCD depends on the amount of transverse momenta with respect to the collision axis ($p_T$) involved in a given scattering, which is closely related to the momentum transfered between the partons (Q). QCD has been very successful in describing quark, antiquark and gluon scatterings involving large amounts of transverse momenta ($Q^2$ above about 10 $GeV^2$)[1)] known as hard interactions. However high-energy pp collisions are dominated by soft parton interactions, so-called minimum bias events. Soft parton interactions also occur in the remains of hard scattering events not associated with the hard process and this is important for many physics analyses such as W boson and top-quark mass measurements [3].

Most of the particles produced at the LHC originate from soft interactions. This is particularly relevant for predictions of background levels associated to many physics processes and also for understanding the complex nature of the radiation environment in which the LHCs detector systems operate.

[1)] We use natural units throughout the thesis: $\hbar = c = 1$.

The total pp cross-section [4] can be divided into elastic and inelastic components, and the inelastic component can be further divided into: non-diffractive, single-diffractive and double-diffractive components.

The total cross-section ($\sigma_{tot}$) can then be written as:

$$\sigma_{tot} = \sigma_{elas} + \sigma_{sd} + \sigma_{dd} + \sigma_{nd} \tag{1.1}$$

where these cross-sections are elastic ($\sigma_{elas}$), single-diffractive ($\sigma_{sd}$), double-diffractive ($\sigma_{dd}$) and non-diffractive ($\sigma_{nd}$), respectively.

Most of the previous measurements of charged particle multiplicity were obtained by selecting data with a double-arm coincidence trigger [5], thus removing large fractions of diffractive events. The data were then further corrected to remove the remaining single-diffractive component. This selection is referred to as non single diffractive (NSD):

$$\sigma_{nsd} = \sigma_{mb} = \sigma_{tot} - \sigma_{elas} - \sigma_{sd} = \sigma_{dd} + \sigma_{nd} \tag{1.2}$$

In some cases [6], designated as inelastic non-diffractive, the residual double-diffractive component was also subtracted. The selection of NSD or inelastic non-diffractive charged particle spectra involves model dependent corrections for the diffractive components and for effects of the trigger selection on events with no charged particles within the acceptance of the detector. The measurement presented in this thesis implements different strategy, which uses a single-arm trigger overlapping with the acceptance of the available tracking volume. **Our minimum bias sample consists of all of the inelastic collisions including single and double diffractive and non-diffractive interactions. Results are presented as inclusive inelastic distributions, with minimal model dependence, by requiring one charged particle within the acceptance of the measurement.**

## 1.2 Phenomenological models of soft hadronic interactions

Low-$p_T$ scattering processes may be described by lowest-order perturbative QCD two-to-two parton scatters, where the divergence of the cross-section at $p_T = 0$ is regulated by phenomenological models. These models include multiple-parton scattering, partonic matter distributions, scattering between the unresolved protons and colour reconnection [7]. The PYTHIA [8] MC event generator implements several of these models.

The parameters of these models have been tuned to describe charged-hadron production and the UE in pp and $p\bar{p}$ data at centre-of-mass energies between 200 GeV and 1.96 TeV [6]. Samples of ten million MC events are produced for single-diffractive, double-diffractive and non-diffractive processes using the PYTHIA 6.4.21 generator.

A specific set of optimised parameters, is the reference **tune** throughout this analysis. These parameters are derived by tuning to UE and minimum bias data from Tevatron at 630 GeV and 1.8 TeV [6]. The MC samples generated with this tune are used to determine detector acceptances and efficiencies and to correct the data.

For the purpose of comparing the present measurement to different phenomenological models describing minimum bias events, the following additional MC samples were generated:

- the ATLAS MC09c [9] PYTHIA tune, which is an extension of the ATLAS MC09 tune optimising the strength of the colour reconnection to describe the $\langle p_T \rangle$ distributions as a function of $n_{ch}$, as measured by CDF in $p\bar{p}$ collisions [6];
- the Perugia0 [10] PYTHIA tune, in which the soft-QCD part is tuned using only minimum bias data from the Tevatron and CERN $p\bar{p}$ colliders;
- the DW [11] PYTHIA tune, which uses the virtuality-ordered showers and is derived to describe the CDF Run II UE and Drell-Yan data;
- finally, the PHOJET generator [12] is used as an alternative model. It describes low-$p_T$ physics using the two-component Dual Parton Model, which includes soft hadronic processes described by Pomeron exchange and semi-hard processes described by perturbative parton scattering.

The non-diffractive, single-diffractive and double-diffractive contributions in the generated samples are mixed according to the generator cross-sections to fully describe the inelastic scattering. All the events are processed through the ATLAS detector simulation program [13], which is based on Geant4 [14]. They are then reconstructed and analysed by the same program chain used for the data. Particular attention is devoted to the description in the simulation of the size and position of the collision beam-spot and of the detailed detector conditions during the data taking.

## 1.3 Data recorded by the ATLAS detector

All data recorded during the stable LHC running periods in December 2009, are used for the analysis at 900 GeV centre-of-mass energy. A total of 455.593 events are collected from colliding proton bunches. A sample of 369.673 events is used for the 7 TeV centre-of-mass energy analysis. In both analyses we require the inner detector (ID) fully operational and the solenoid magnet on.

After the offline selection cuts explained in the Chapters 4 and 5, a total of 326.201 and 369.673 events are kept at 900 GeV and 7 TeV respectively, which contains a total of 1.863.622 and 3.769.168 selected tracks. We analyze approximately the same number of events but we have twice more tracks at 7 TeV because the increased centre-of-mass energy leads to a larger number of charged particles per event.

At 900 GeV the integrated luminosity for the final event sample is estimated using a sample of events with energy deposits in both sides of the forward and end-cap calorimeters. The MC based efficiency and the PYTHIA default cross-section of 52.5 mb are then used to determine the luminosity of the data sample to be approximately $9\mu b^{-1}$, while the maximum instantaneous luminosity is approximately $5 \times 10^{26} cm^{-2} s^{-1}$.

At 7 TeV the integrated luminosity is about 6.8 $\mu b^{-1}$. More details are given in the Section 4.5 dedicated to the luminosity measurements in ATLAS.

We also have data recorded at 2.36 TeV but the LHC was not providing stable beams. ATLAS does not use full ID with unstable LHC beams. For this reason we do not include 2.36 TeV data in our studies.

# Chapter 2

# The ATLAS detector

The LHC at CERN is able to extend the frontiers of particle physics with its unprecedented high-energy and luminosity. Inside the LHC, bunches of up to $10^{11}$ protons will collide 40 million times per second to provide 14 TeV pp collisions at a design luminosity of $10^{34} cm^{-2} s^{-1}$. At the moment we run at 7 TeV while we have reached the nominal numbers of protons per bunch. The peak of stable luminosity delivered in 2010 is $2.07 \times 10^{32} cm^{-2} s^{-1}$.

The high interaction rates, radiation doses, particle multiplicities and energies, as well as the requirements for precision measurements have set new standards for the design of particle detectors. Two general purpose detectors, ATLAS and CMS (Compact Muon Solenoid) have been built for probing pp collisions. This chapter presents a comprehensive overview of the ATLAS detector during the first LHC collisions. This detector represents the work of a large collaboration of several thousand physicists, engineers, technicians, and students over a period of fifteen years of dedicated design, development, fabrication, and installation.

The high luminosity and increased cross-sections at the LHC enable further high precision tests of QCD, electroweak interactions, and flavour physics. The top quark will be produced at the LHC at a rate of a few tens of Hz, providing the opportunity to test its couplings and spin.

The search for the Standard Model Higgs boson has been used as a benchmark to establish the performance of important sub-systems of ATLAS. It is a particularly important process since there is a range of production and decay mechanisms, depending on the mass of the Higgs boson (H). At low masses ($m_H < 2m_Z$), the natural width would only be a few MeV, and the observed width would be defined by the instrumental resolution. The predominant decay mode into hadrons would be difficult to detect due to QCD backgrounds, and the two-photon decay channel would be an important one. Other promising channels could be, for example, associated production of H such as $t\bar{t}$H, WH, and ZH, with $H \to b\bar{b}$, using a lepton from the decay of one of the top quarks or of the vector boson for triggering and background rejection. For masses above 130 GeV, Higgs-boson decays, $H \to ZZ^*$, where each Z decays to a pair of oppositely charged leptons, would provide the experimentally cleanest channel to study the properties of the Higgs boson. For masses above approximately 600 GeV, WW and ZZ decays into jets or involving neutrinos would be needed to extract a signal. The tagging of forward jets from the WW or ZZ fusion production mechanism are important for the discovery of the Higgs boson.

The formidable LHC luminosity and resulting interaction rate are needed because of the small cross-sections expected for many of the processes mentioned above. However, with an inelastic pp cross-section of 80 mb (at 14 TeV), the LHC will produce a total rate of $10^9$ inelastic events/s at design luminosity of $10^{34} cm^{-2} s^{-1}$. This presents a serious experimental difficulty as it implies that every candidate event for new physics will on the average be accompanied by twenty three inelastic events per bunch crossing.

The nature of pp collisions imposes another difficulty. QCD jet production cross-sections dominate over the rare processes mentioned above, requiring the identification of experimental signatures characteristic of the physics processes in question, such as $E_T^{miss}$ or secondary vertices. Identifying such final states for these rare processes imposes further demands on the integrated luminosity needed, and on the particle identification capabilities of the detector.

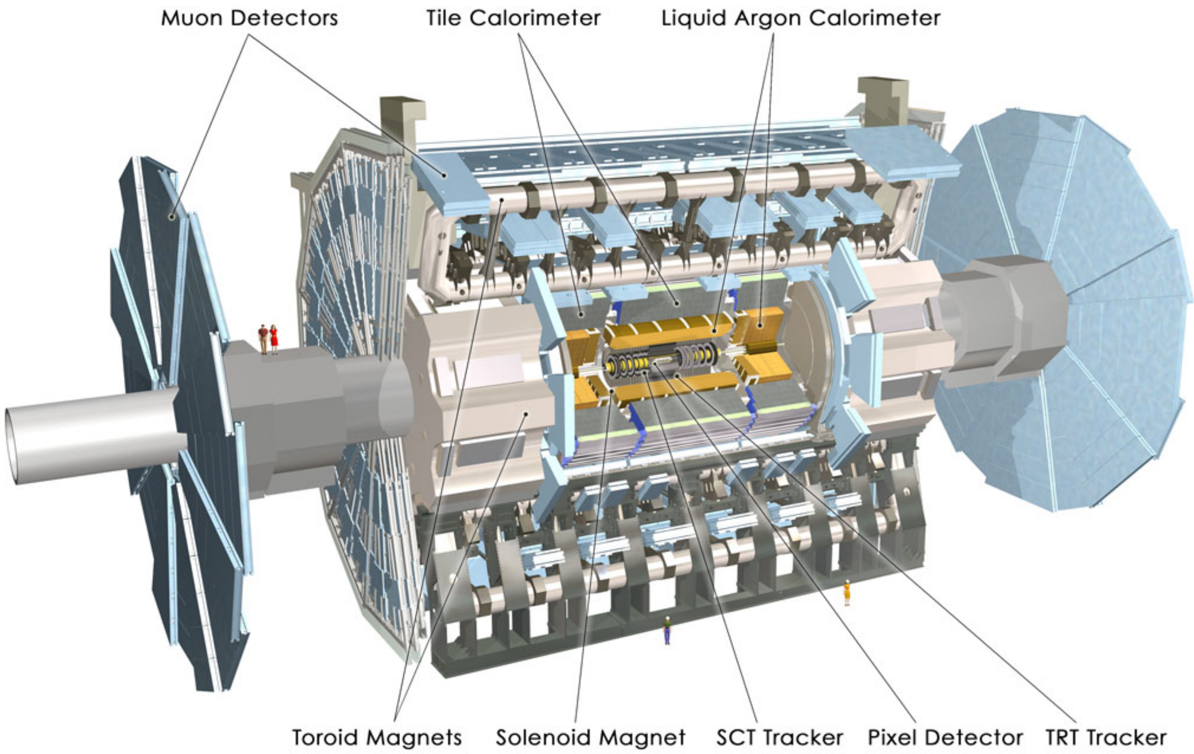



Figure 2.1: The ATLAS detector.

The overall ATLAS detector layout is shown in Figure 2.1 and its main performance goals are listed in Table 2.1.

The ATLAS detector is nominally forward-backward symmetric with respect to the interaction point[1]. The magnet configuration comprises a thin superconducting solenoid

[1] The ATLAS reference system is a cartesian right-handed coordinate system, with the nominal interaction point at the origin. The anti-clockwise beam direction defines the positive z-axis, while the positive x-axis is defined as pointing from the collision point to the centre of the LHC ring and the positive y-axis points upwards. The azimuthal angle is measured around the beam axis, and the polar angle, $\theta$, is measured with respect to the z-axis. The pseudorapidity is defined as $\eta = -\ln[\tan(\frac{\theta}{2})]$.

| Detector component | Required resolution | $\eta$ coverage | |
|---|---|---|---|
| | | Measurement | Trigger |
| Tracking | $\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$ | $\pm 2.5$ | |
| EM calorimetry | $\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$ | $\pm 3.2$ | $\pm 2.5$ |
| Hadronic calorimetry (jets) | | | |
| barrel and end-cap | $\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$ | $\pm 3.2$ | $\pm 3.2$ |
| forward | $\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$ | $3.1 < \lvert\eta\rvert < 4.9$ | $3.1 < \lvert\eta\rvert < 4.9$ |
| Muon spectrometer | $\sigma_{p_T}/p_T$=10% at $p_T$ = 1 TeV | $\pm 2.7$ | $\pm 2.4$ |

Table 2.1: General performance goals of the ATLAS detector. The units for E and $p_T$ are in GeV. The fractional energy resolution is conventionally parametrised as $\sigma(E)/E = a/E \oplus b/\sqrt{E} \oplus c$, where a is the noise term, b is the sampling term and c is the constant term.

surrounding the ID cavity, and three large superconducting toroids (one barrel and two end-caps) arranged with an eight-fold azimuthal symmetry around the calorimeters. The ID is immersed in a 2 T solenoidal field.

Pattern recognition, momentum and vertex measurements, are achieved with a combination of discrete, high-resolution semiconductor pixel and strip detectors in the inner part of the tracking volume, and straw-tube tracking detectors with the capability to generate and detect transition radiation in its outer part [15].

High granularity liquid-argon (LAr) electromagnetic sampling calorimeters, with excellent performance in terms of energy and position resolution, cover the pseudorapidity range $|\eta| < 3.2$. The hadronic calorimetry in the range $|\eta| < 1.7$ is provided by a scintillator-tile calorimeter, which is separated into a large barrel and two smaller extended barrel cylinders, one on either side of the central barrel. In the end-caps ($|\eta| > 1.5$), LAr technology is also used for the hadronic calorimeters, matching the outer $|\eta|$ limits of end-cap electromagnetic calorimeters. The LAr forward calorimeters provide both electromagnetic and hadronic energy measurements, and extend the pseudorapidity coverage to $|\eta| = 4.9$.

The calorimeter is surrounded by the muon spectrometer. The air-core toroid system, with a long barrel and two inserted end-cap magnets, generates strong bending power in a large volume within a light and open structure. Multiple scattering effects are thereby minimised, and excellent muon momentum resolution is achieved with three layers of high precision tracking chambers.

The pp interaction rate at the design luminosity of $10^{34} cm^{-2} s^{-1}$ is approximately 1GHz, while the event data recording, based on technology and resource limitations, is limited to about 300 Hz. This requires an overall rejection factor of $5 \times 10^6$ against minimum bias processes while maintaining maximum efficiency for the new physics. The Level-1 (L1) trigger system uses a subset of the total detector information to make a decision on whether or not to continue processing an event, reducing the data rate to approximately 30 kHz (limited by the bandwidth of the readout system, which is upgradeable to 100 kHz). The subsequent two levels, collectively known as the high-level trigger, are the Level-2 (L2) trigger and the event filter. They provide the reduction to a final data taking rate of approximately 300 Hz.

In the following we describe in more details the detector components that are relevant in this analysis: the tracking system, the calorimeters, the solenoidal magnetic field and the trigger system.

## 2.1 The tracking devices

The pseudorapidity and transverse momentum distributions of charged particles are measured using the ATLAS ID. The ATLAS ID is described in detail elsewhere [15].

The ATLAS ID tracker, shown in Figure 2.2, is made out of three detector types, moving inside out we find: the silicon pixel detector the SemiConductor Tracker (SCT), and the Transition Radiation Tracker (TRT). All these detectors allow precision measurement of charged particle trajectories in the environment of numerous tracks, but the pixel mainly contributes to the accurate measurement of the vertices. The SCT helps to measure precisely the particle momenta and the TRT eases the pattern recognition with its very large number of close hits (while also contributing to electron identification). Figure 2.3 illustrates the layout and shows that each detector consists of barrel and endcap regions. The ID is mounted inside a solenoid magnet which provides a 2 T magnetic field generated by the central solenoid, which extends over a length of 5.3 m with a diameter of 2.5 m.

The precision tracking detectors (pixel and microstrip) cover $|\eta| < 2.5$ and are divided into barrel ($|\eta| < 1.4$) and endcaps ($1.4 < |\eta| < 2.5$). In the barrel region, they are arranged on concentric cylinders around the beam axis while in the end-cap regions, they are located on disks perpendicular to the beam axis. The highest granularity is achieved around the vertex region using silicon pixel sensors.

The ID sub-detectors are designed as independent but also complementary systems. The radius and lengths of the sub-detctors are summarized in the Table 2.2.

The pixel layers are segmented in $R-\phi$ and z with typically three pixel layers crossed by each track. The first layer, called the vertexing layer, is at a radius of 51 mm. The designed intrinsic accuracies in the barrel are 10 $\mu$m (R-$\phi$) and 115 $\mu$m (z) and in the disks are 10 $\mu$ (R-$\phi$) and 115 $\mu$m (R). The pixel detector has approximately 80.4 million readout channels.

For the SCT, eight strip layers (four space points) are crossed by each track. In the barrel region, this detector uses small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring R-$\phi$. Each side of a detector module consists of two 6.4 cm long, daisy chained sensors with a strip pitch of 80 $\mu$m. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately 80 $\mu$m. The intrinsic accuracies per module in the barrel are 17 $\mu$m (R-$\phi$) and 580 $\mu$m (z) and in the disks are 17 $\mu$m (R-$\phi$) and 580 $\mu$m (R). The total number of readout channels in the SCT is approximately 6.3 million.

A large number of hits (typically 30 per track, see Figure 5.6) is provided by the 4 mm diameter straw tubes of the TRT, which enables track following up to $|\eta| = 2.0$. The TRT only provides R-$\phi$ information, for which it has an intrinsic accuracy of 130 $\mu$m per

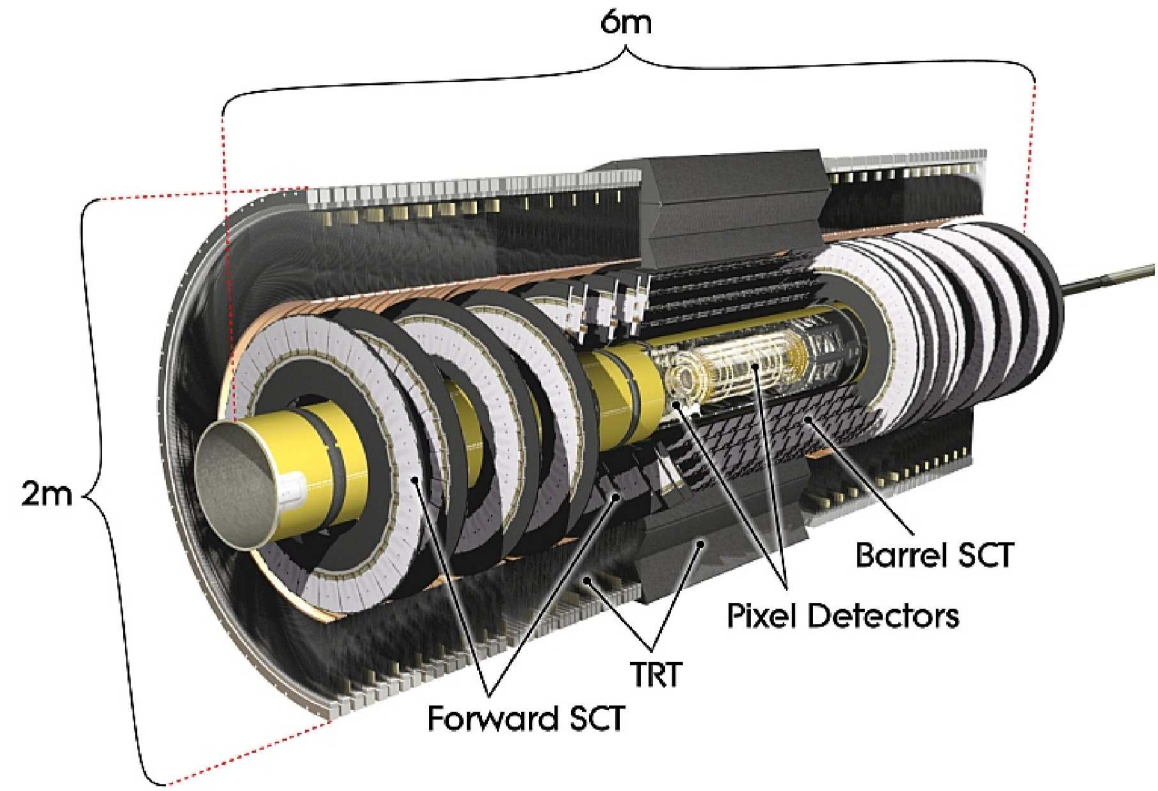



Figure 2.2: Cut-away view of the ATLAS ID.

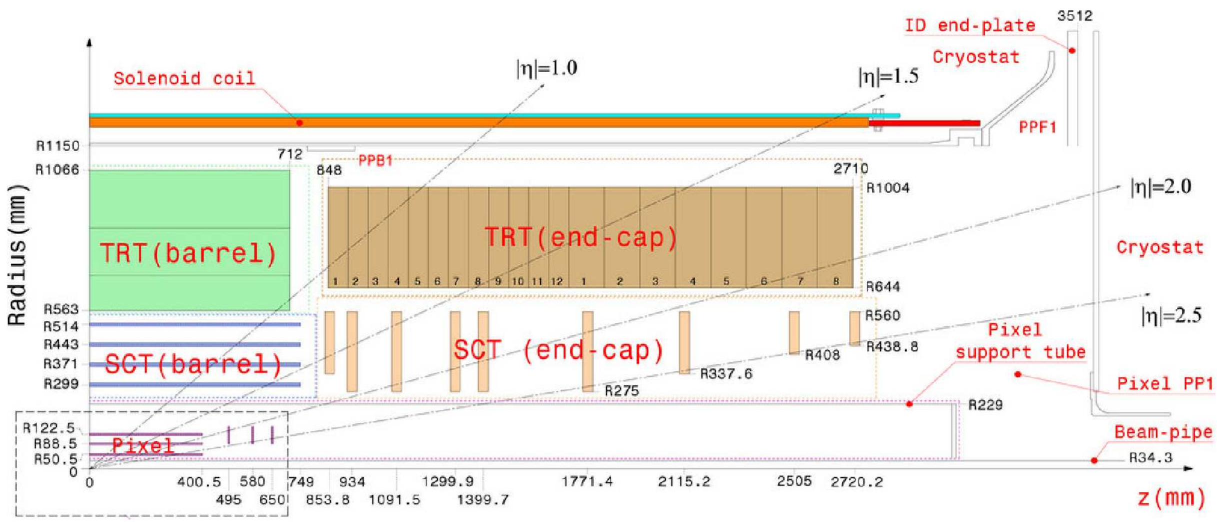



Figure 2.3: Plan view of a quarter section of the ATLAS ID showing each of the major elements with its active dimensions.

straw. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $|\eta| = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The total number of TRT readout channels is approximately 351.000.

| | | Radius (mm) | Length (mm) |
|---|---|---|---|
| **Pixel** | Barrel (3 layers) | $50.5 < R < 122.5$ | $0 < \|z\| < 400.5$ |
| | End-cap (2x3 disks) | $88.8 < R < 149.6$ | $495 < \|z\| < 650.0$ |
| **SCT** | Barrel (4 layers) | $299 < R < 514$ | $0 < \|z\| < 749$ |
| | End-cap (2x9 layers) | $275 < R < 560$ | $839 < \|z\| < 2735$ |
| **TRT** | Barrel (73 straw planes) | $563 < R < 1066$ | $0 < \|z\| < 712$ |
| | End-cap (160 straw planes) | $644 < R < 1004$ | $848 < \|z\| < 2710$ |

Table 2.2: Radius and lengths of the ID sub-detctors.

The ATLAS ID provides hermetic and robust pattern recognition, excellent momentum resolution and both primary and secondary vertex measurements for charged tracks above a $p_T$ threshold, which is nominally 500 MeV but can be as low as 100 MeV within $|\eta| < 2.5$. The track reconstruction efficiency at low momentum is limited because of the large material effect in the ID.

A $p_T$ cut of 500 MeV, used in this analysis, corresponds to tracks that traverse the precision Si tracker (pixels and SCT) allowing low-$p_T$ tracks to be well reconstructed.

### 2.1.1 Track reconstruction

Track reconstruction begins when silicon clusters are formed from raw hits. Next, the three dimensional space points are formed from these clusters. Pixel clusters translate directly into space points, however SCT clusters from each side of a module must be associated to form a single SCT space point.

Several different pattern recognition algorithms are used to find tracks in the ID. The tracks typically used in physics analyses are found using the inside-out pattern recognition algorithm [15], which starts close to the interaction point with silicon space points and extends outwards to the TRT. Track seeds are created from three space points and then used to define roads to search for clusters or single sided SCT hits to associate to the track. The next step, ambiguity processing, refines the tracks using a more sophisticated track fitter and removes overlapping tracks. Lastly, these silicon tracks are used to define roads in the TRT in which TRT extension can be attached.

The inside-out sequence is followed by an outside-in (also referred to as back tracking) sequence that runs on the remaining hit collection after hits that are already part of tracks have been removed. The back tracking algorithm is seeded in the TRT and the track is extrapolated into the silicon. A second inside-out sequence with less stringent pattern recognition requirement is executed afterwards.

Other algorithms include low $p_T$ tracking and standalone TRT tracking. The low $p_T$ tracking is optimized to find tracks with $p_T < 500$ MeV. It is ran after the standard inside-out pattern recognition and uses seeds formed out of pixel clusters, building tracks from clusters which have not been associated to inside-out tracks.

The low $p_T$ algorithm will be used for the second paper of minimum bias at 7 TeV. In the analysis presented in this thesis we only use the inside-out algorithm to find tracks with $p_T > 500$ MeV. The procedure how the tracking efficiency is determined in MC is described in Section 6.3.1.

## 2.2 The principles of the Tile Calorimeter

Electromagnetic and hadronic calorimeters are indispensable components of a general purpose hadron collider detector. Jointly they must provide accurate energy and position measurements of electrons, photons, isolated hadrons, jets, and transverse missing energy, as well as helping in particle identification and in muon momentum reconstruction. The electromagnetic (EM) and hadronic compartments of the ATLAS calorimeter system cover the pseudorapidity region $|\eta| < 4.9$. The EM compartments are LAr sampling calorimeters[2)], while the detector media of the hadronic calorimeters differ according to the $\eta$ region. The Tile Calorimeter (TileCal) [16] is a sampling plastic scintillator/iron detector, located in the region $|\eta| < 1.7$; it is divided into three cylindrical sections, referred to as the central barrel and extended barrels. The barrel covers the region $-1.0 < \eta < 1.0$, and the extended barrels cover the region $0.8 < |\eta| < 1.7$. Radially, the TileCal extends from an inner radius of 2.28 m to an outer radius of 4.25 m. Before reaching TileCal, particles from the collision region first traverse the inner tracking detectors and the barrel or endcap sections of the LAr/lead EM calorimeters[3)].

The TileCal is a sampling device made out of steel and scintillating tiles, as absorber and active material respectively. It realizes a simple and very well proven idea of calorimetry, particularly suited for the LHC environment. The absorber structure is a laminate of steel plates of various dimensions, connected to a massive structural element referred to as a girder, which is placed at the outer part of the calorimeter. The TileCal scintillator plates are placed perpendicular to the colliding beam axis, and are also radially staggered in depth. The structure is periodic along the beam axis. The tiles are 3 mm thick and the total thickness of the iron plates in one 18 mm period is 14 mm. The highly periodic structure of the system allowed the construction of a large detector by assembling relatively small sub-modules together. Since the mechanical assembly was completely independent from the optical instrumentation, the design became simple and cost effective. Simplicity had been the guideline for the light collection scheme used as well: fibres are coupled radially to the tiles along the outside faces of each module[4)]. The laminated structure of the absorber allows for grooves in which the wave-length shifting (WLS) fibres run. The use of fibre readout allowed to define a 3D TileCal cell geometry by arranging the fiber readout in space and creating quasi-projective tower structure, where the deviations from perfect projectivity are small compared to the typical angular extent of hadronic jets. By the grouping of WLS fibers to specific photon detectors (in our case photomultipliers,

[2)] In sampling calorimeters the functions of particle absorption and active signal readout are separated. This allows to optimize the choice of absorber material.

[3)] For $|\eta| > 1.5$, particles also traverse the LAr/copper end-cap calorimeter.

[4)] Tile module corresponds to 1/64 slice in $\phi$ of any given calorimeter cylinder. One module is equipped with a complete set of the readout electronics and can be operated independently from the rest of the calorimeter.

PMTs), modules are segmented in $\eta$ and in radial depth (1.5, 4.1 and 1.8 $\lambda_{int}$ in the barrel; 1.5, 2.6, 3.3 $\lambda_{int}$ in the extended barrels), resulting in typical cell dimensions of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ ($0.1 \times 0.2$ in the last depth). Altogether, TileCal comprises 4.672 readout cells, each equipped with two PMTs that receive light from opposite sides of every tile.

A compact electronics readout is housed in the girder of each module. The readout of the two sides of each of the scintillating tiles into two separate PMTs guarantees a sufficient light yield. For readout we use the R7877 Hamamatsu photomultiplier[5)]. Ionizing particles crossing the tiles induce the production of light in the base material of the tiles, with wavelength in the UV range which subsequently is converted to visible light by scintillator dyes. This scintillation light propagates through the tile to its edges, where it is absorbed by the WLS fibres and shifted to a longer wavelength chosen to match the sensitive region of the PMT photocathode. A fraction of the light re-emitted in the fibre is captured and propagated via total internal reflection to the PMT where it is converted into electrical signal. A light mixer is placed between the fibres and the photocathode to optimize uniformity of the photon collection. The ends of the fibres away from the PMTs are aluminized to increase amount of the light captured and delivered to PMTs.

Part of the gap between calorimeter cylinders contains a structural extension of the extended barrel: the Intermediate Tile Calorimeter (ITC), which is a calorimeter extension in order to maximize the volume of active material in the gap region, while still leaving room for the services and cables, which mainly belong to the tracker. The ITC consists of a calorimeter plug between the region $0.8 < |\eta| < 1.0$, and, due to severe space constraints, bare scintillator between $1.0 < |\eta| < 1.6$. The scintillators in the region $1.0 < |\eta| < 1.2$ are called gap scintillators, and the scintillators between $1.2 < |\eta| < 1.6$ are called crack scintillators.

### 2.2.1 The Minimum Bias Trigger Scintillators (MBTS)

Barrel modules have two fingers[6)] connected from the left and the right sides. The extended barrel modules have only one finger at one side and ITC/gap/crack/MBTS connected on the side facing center of ATLAS. In the Figure 2.4 we can see components in the extended barrel module. More details on the structure of the calorimeter volume can be found in [16].

The gap and crack scintillators supplement the TileCal calorimeter by attempting to recover for energy lost in the inactive material in the gap between the TileCal barrel and extended barrel calorimeters and in the crack between the LAr barrel and end cap calorimeters respectively.

The MBTS are designed to function only during initial data taking at low luminosities. After some months of higher luminosity operation the scintillators will yellow due to radiation damage. The two sets of sixteen scintillator counters are installed on the inner face of the end-cap calorimeter cryostats and are used to trigger on minimum bias events.

5) These PMTs were extensively tested in Pisa, where I made a PMT dedicated studies resulting in 'Laurea' defence [17].

6) In the gap region between barrel and extended barrel the outer girder of the tile is made of fingers and holes in between for cable passages.

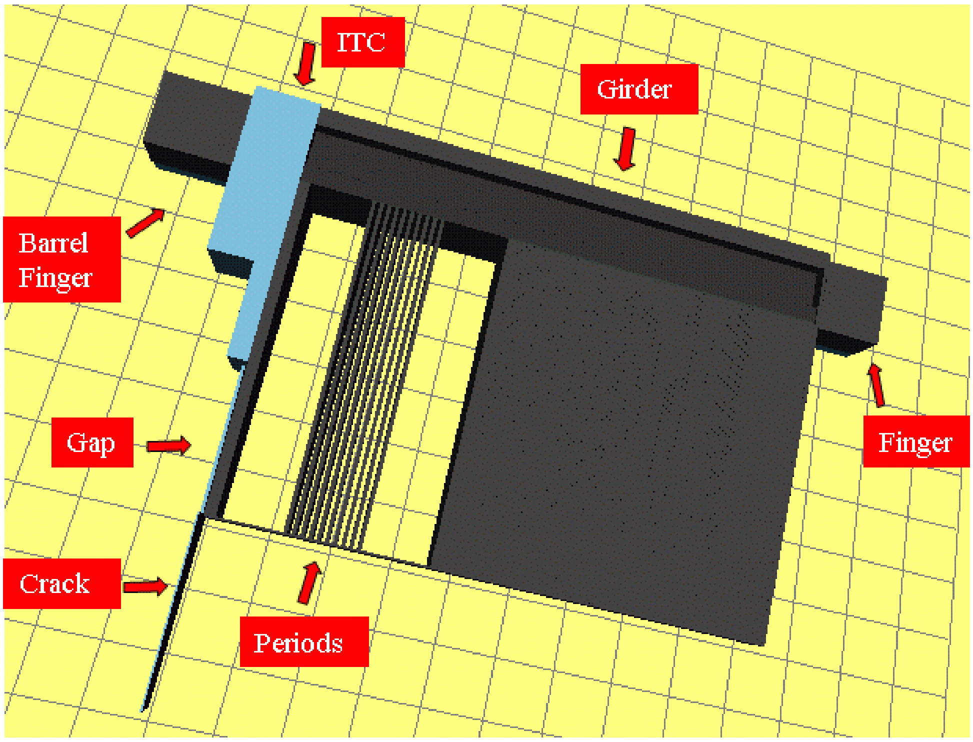


Figure 2.4: Components in the extended barrel module of TileCal.

Each set of counters is segmented in eight units in $\phi$ and two units in $\eta$ (see Figures 2.5 and 2.6). They are located at $|z| = 3.560$ $mm$, the innermost set covers radii between 153 $mm$ and 426 $mm$, corresponding to the region $2.82 < |\eta| < 3.84$ and the outermost set covers radii between 426 $mm$ and 890 $mm$, corresponding to the region $2.09 < |\eta| < 2.82$.

For the TileCal cells the scintillation light is collected by the WLS fibres that are routed to a PMT. These cells (gap/crack/MBTS) are not part of the iron-scintillator sampling structure of TileCal. They are simple scintillators read out by a single PMT.

The TileCal PMTs are connected to the standard TileCal 3in1 cards. The 3in1 cards provides three outputs: a trigger signal, readout at low gain and readout at high gain. There is one important difference for the MBTS readout. In order to obtain a better signal to noise ratio in the trigger signal, the TileCal 3in1 cards there were modified to send the high gain output to the trigger, not the low gain output that is set in all other TileCal channel.

In conclusion, the MBTS is read out through the TileCal electronics providing a fast L1 signal, which is discriminated above a voltage threshold, relative to the bunch-crossing signal. More details will be given in the section dedicated to the minimum bias trigger (Section 2.4).

## 2.3 The solenoidal magnetic filed

This section is dedicated to a brief description of the ATLAS magnet system, which consists of :

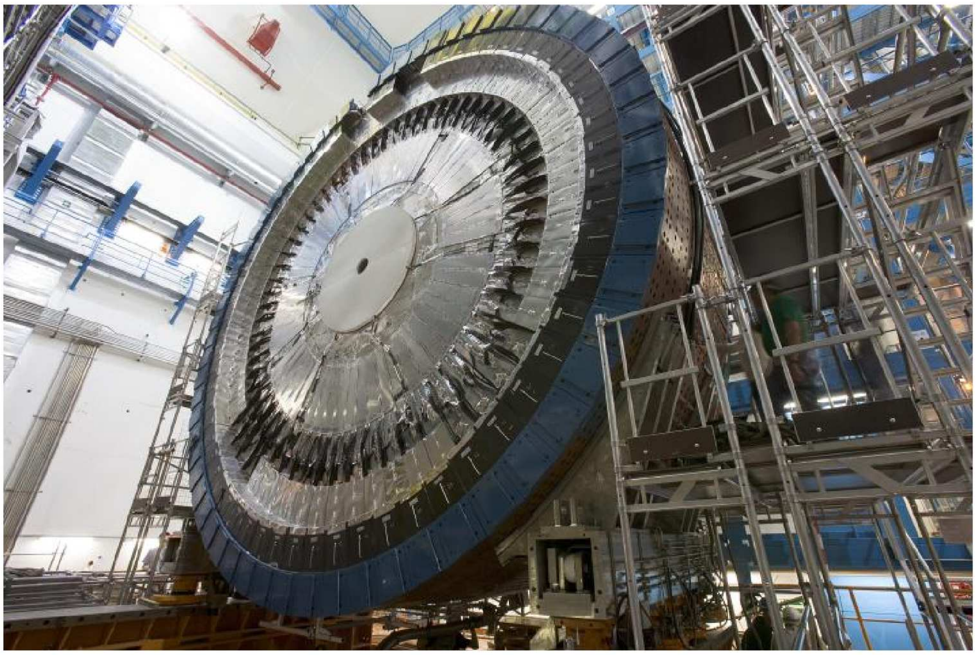


Figure 2.5: Picture of the inner surface of the LAr endcap cryostats where the MBTS are mounted.

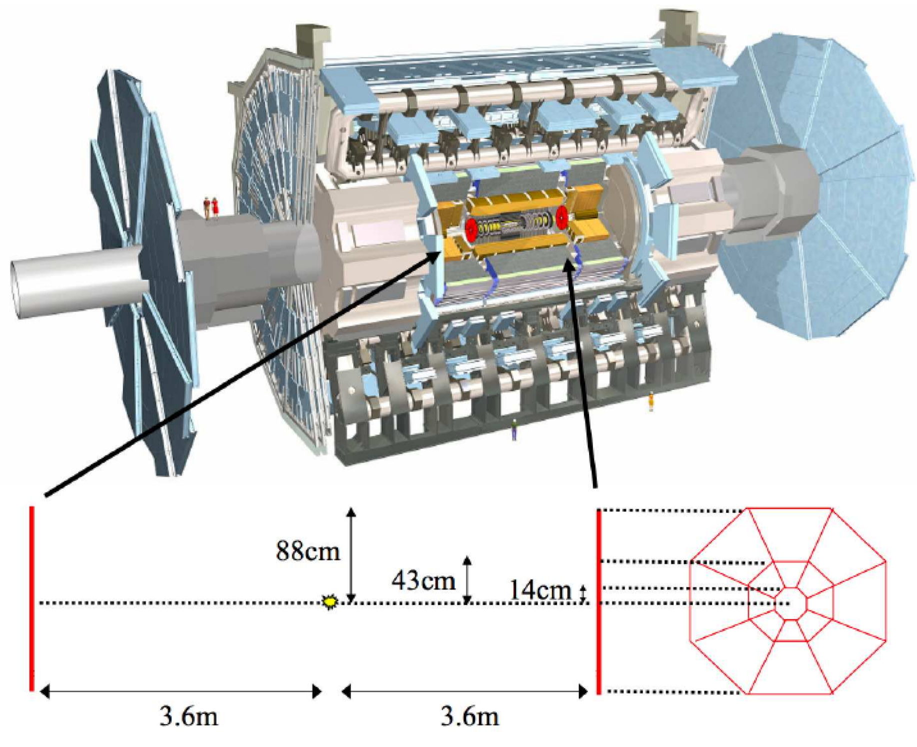



Figure 2.6: Geometry of 16 MBTS counters. The dimensions and the position relative to ATLAS detector are also shown.

- a solenoid, which is aligned on the beam axis and provides a 2 T axial magnetic field for the ID, while minimising the radiative thickness in front of the barrel electromagnetic calorimeter;
- a barrel toroid and two end-cap toroids, which produce a toroidal magnetic field of approximately 0.5 T and 1 T for the muon detectors in the central and end-cap regions, respectively.

The spatial arrangement of the coil windings is shown in Figure 2.7. The central solenoid is designed to provide a 2 T axial field (1.998 T at the magnets centre at the nominal 7.730 kA operational current). To achieve the desired calorimeter performance, the layout is carefully optimised to keep the material thickness in front of the calorimeter as low as possible, resulting in the solenoid assembly contributing a total of 0.66 radiation lengths at normal incidence.

This required, in particular, that the solenoid windings and LAr calorimeter share a common vacuum vessel, thereby eliminating two vacuum walls. The inner and outer diameters of the solenoid are 2.46 m and 2.56 m and its axial length is 5.8 m. The coil mass is 5.4 tonnes and the stored energy is 40 MJ. The stored energy to mass ratio of only 7.4 kJ/kg at nominal field clearly demonstrates successful compliance with the design requirement of an extremely light weight structure.

The flux is returned by the steel of the ATLAS hadronic calorimeter and its girder structure. The solenoid is charged and discharged in about 30 minutes. In the case of a quench, the stored energy is absorbed by the enthalpy of the cold mass which raises the cold mass temperature to a safe value of 120 K maximum. Re-cooling to 4.5 K is achieved within one day.

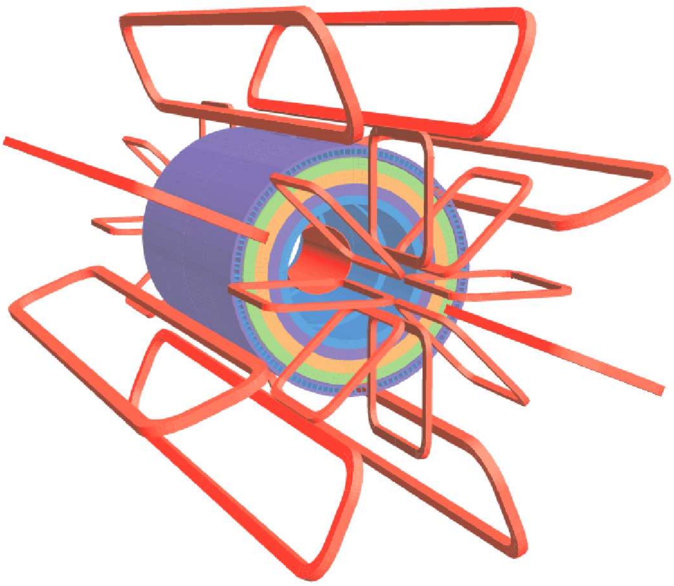


Figure 2.7: Geometry of magnet windings and TileCal steel. The eight barrel toroid coils, with the end-cap coils interleaved are visible. The solenoid winding lies inside the calorimeter volume. The TileCal calorimeter is modeled by four layers with different magnetic properties, plus an outside return yoke.

## 2.4 The trigger system

As described before, the trigger consists of three levels of event selection: L1, L2, and event filter. The L2 and event filter together form the High-Level Trigger (HLT).

The L1 trigger is implemented using custom made electronics, while the HLT is almost entirely based on commercially available computers and networking hardware. The L1 trigger searches for signatures from high-$p_T$ muons, electrons/photons, jets, and $\tau$-leptons decaying into hadrons. It also selects events with large missing transverse energy ($E_T^{miss}$) and large total transverse energy. The L1 trigger uses reduced-granularity information from a subset of detectors: the Resistive Plate Chambers and Thin-Gap Chambers for high-$p_T$ muons, and all the calorimeter sub-systems for electromagnetic clusters, jets, $\tau$-leptons, $E_T^{miss}$, and large total transverse energy. The maximum L1 accept rate which the detector readout systems can handle is 75 kHz (upgradeable to 100 kHz), and the L1 decision must reach the front-end electronics within 2.5 $\mu s$ after the bunch-crossing with which it is associated.

The L2 trigger is seeded by Regions-of-Interest (RoI‘s). These are regions of the detector where the L1 trigger has identified possible trigger objects within the event. The L2 trigger uses RoI information on coordinates, energy, and type of signatures to limit the amount of data which must be transferred from the detector readout. The L2 trigger reduces the event rate to below 3.5 kHz, with an average event processing time of approximately 40 ms.

The event filter uses offline analysis procedures on fully built events to further select events down to a rate which can be recorded for subsequent offline analysis. It reduces the event rate to approximately 200 Hz, with an average event processing time of order of four seconds.

The HLT algorithms use the full granularity and precision of calorimeter and muon chamber data, as well as the data from the ID, to refine the trigger selections. Better information on energy deposition improves the threshold cuts, while track reconstruction in the ID significantly enhances the particle identification (for example, distinguishing between electrons and photons). The event selection at both L1 and L2, primarily uses inclusive criteria, for example, high-$E_T$ objects above defined thresholds.

The data acquisition system receives and buffers the event data from the detector specific readout electronics at the L1 trigger rate. The data transmission is performed over point-to-point Readout Links (ROL‘s). It transmits to the L2 trigger any data requested by the trigger (typically the data corresponding to RoIs) and, for those events fulfilling the L2 selection criteria, event-building is performed. The assembled events are then moved by the data acquisition system to the event filter, and the events selected there are moved to permanent event storage.

In addition to controlling movement of data down the trigger selection chain, the data acquisition system also provides for the configuration, control and monitoring of the ATLAS detector during data taking. Supervision of the detector hardware (gas systems, power-supply voltages, etc.) is provided by the Detector Control System.

### 2.4.1 Minimum bias trigger scenarios

A minimum bias trigger should select inelastic collisions with as little bias as possible, precluding the use of the standard high-$p_T$ triggers. Ideally, the L1 random trigger with beam pickup would be used to accept events with zero bias, and inelastic collisions would be selected offline. However, during early running when the luminosity below $10^{30} cm^{-2} s^{-1}$, the random trigger was inefficient since the probability of an interaction during a bunch crossing is $< 1\%$.

The events used to make the minimum bias measurements are collected with the MBTS. The distributions must therefore be corrected for trigger inefficiencies that can potentially cause a bias. Section 6.1 describes the measurement of the trigger efficiency for this analysis. Note that the efficiency is measured with respect to the offline selection criteria based on ID tracks (Chapter 5) and therefore can only be applied to events with the same selection criteria.

In this study events from the beam pickup based timing system (BPTX) stream are used also, with the reconstruction, skimming and Bunch Crossing ID cuts, and data quality requirements described in Reference [18].

ATLAS uses several independent detector components for the selection of inelastic pp interactions with minimum bias. In the following we describe the triggers used for the selection of the events and those used to estimate the trigger efficiency.

#### The ATLAS beam pickup based timing system

The ATLAS BPTX detectors [19] utilize the same electrostatic beam pickup system that the LHC uses to monitor the transverse position of the beam inside the beam pipe. The BPTX beam pickup stations enable ATLAS to ‘see’ the bunches of particles on their way into the center of the experiment. The BPTX stations are located along the LHC on both sides of ATLAS, 175 m away from the interaction point. The signals from the BPTX stations are transmitted over 200 m low-loss cable into the ATLAS underground counting room.

The BPTX signals are fed into a discriminator and then into the Central Trigger Processor (CTP) to provide filled bunch triggers.

The BPTX pulses are discriminated and shaped into 25 ns long pulses and used as input to the CTP. This filled bunch trigger enables the trigger system to know at which clock ticks there are particle bunches in the interaction region from either beam.

These trigger signals are used to form a prescaled L1 trigger, and can be combined with other triggers to ensure there is a beam passing through ATLAS.

#### MBTS readout

The MBTS detector presented in Section 2.2.1 consists of 32 scintillator counters 2 cm tick, organised into 2 disks, one on each detector side of ATLAS (A and C).

The MBTS signals, after being shaped and amplified by the TileCal electronics, are further amplified by a factor of two and fed into leading edge discriminators. The NIM pulses are stretched and sent to the CTP. An MBTS hit is defined as a MBTS signal above a discriminator threshold of 60 mV. At the CTP input these signals are stretched to 200 ns to negate the effect of time drift and the MBTS multiplicity is calculated. The MBTS multiplicity is calculated for each side independently.

From such inputs three relevant L1 trigger items are formed: MBTS_1, MBTS_2 and MBTS_1_1. These items require a BPTX signal, from either side, and respectively at least one MBTS hit, at least two MBTS hits and at least one MBTS hit per side. All three triggers are running unprescaled with no additional requirements in the High-Level Trigger for the entire datasets used in this analysis.

#### Inner detector based minimum bias trigger

The ID minimum bias trigger, called in the following mbSpTrk, provides an alternative method to the MBTS system to select inelastic interactions. MbSpTrk is seeded by the BPTX trigger at L1. The event selection takes place at the HLT. The mbSpTrk trigger uses the ID silicon sub-detectors: the pixel and the silicon central tracker (SCT). The trigger therefore covers the complete $\eta$-region up to $|\eta| < 2.5$ where track based measurements are performed. MbSpTrk stream is shown in Figure 2.8.

At the L2 trigger the algorithms detect central detector activity by forming space points in both the pixel and SCT detectors. The space points are 3-dimensional hit representations formed from hit clusters. The different detector technology of the pixel and SCT systems is reflected in the space point formation respectively. While the pixel space points are made by a direct transformation of pixel clusters into the space points, the SCT space points are only created, if a pair of strip clusters originate from opposite sides of a module and overlap in $\eta$ and $\phi$. This coincidence of silicon strips implies an intrinsic noise suppression in the SCT detector. To have a better handle of electronic noise hits in the pixel detector, a requirement on the pixel cluster time over threshold of twenty bunch crossings is applied before forming the total number of pixel space points in the event. Since the main task at L2 is to suppress empty bunch crossing events,

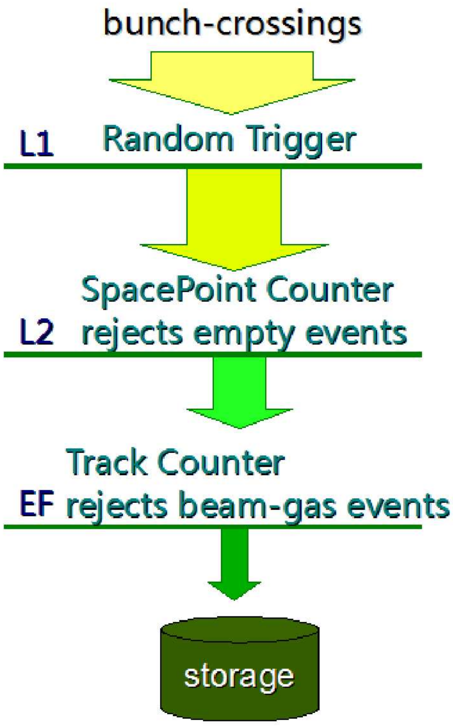



Figure 2.8: MbSpTrk trigger slice. The Event Filter part of the ATLAS trigger, for the moment, accepts all event due to low beam-background rates.

the thresholds are set to operate just above the electronic noise level. For the events considered in the analysis, a total number of seven pixel and seven SCT space points are required.

At the third trigger level, the Event Filter, silicon tracks are reconstructed, employing sequences of the standard tracking where the same reconstruction tools are used as in the offline track reconstruction software. For the ID minimum bias trigger, the minimum reconstruction $p_T$ is lowered to 200 MeV from the default of 500 MeV. The trigger cut is applied on the number of tracks which have a longitudinal impact parameter $z_0$ that is close enough to the nominal beam-spot, i.e. at least one reconstructed silicon track within $|z_0| \leq 200$ mm is needed to record the event. This trigger level has been enabled to remove the beam induced background interactions, beam-gas and beam-halo.

## 2.5 Summary

The minimum bias triggers are provided by the MBTS and the silicon based tracking detectors. Triggering is facilitated by a three level architecture starting with hardware based triggers at L1. The individual sub-detectors send L1 decisions to the CTP which sets the global trigger bits. Software decisions are made at L2 using partial detector information with the exception of the ID, for which all hits are read out. The third level, the Event Filter, does a full event reconstruction in order to select the final events.

The MBTS trigger efficiencies are measured in data using events passing mbSpTrk as the reference data sample, more details are given in the section 6.1 dedicated to the trigger efficiencies.

The goal is to keep the most efficient trigger and to reduce model dependencies in corrections to charged particle distributions.

Since the 2009 data taking period at $\sqrt{s} = 900\ GeV$, several changes have been made to the MBTS detector that is used to trigger events for readout. The counter thresholds and voltages are tuned to get a better signal to noise separation and therefore we observe a slightly higher trigger efficiency during the 2010 data taking period at $\sqrt{s} = 7\ TeV$.

# Chapter 3

# Monte Carlo simulations

This chapter describes the MC studies to be compared with the analysis of minimum bias data taken at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV. A brief descriptions of the tunes (MC09, Perugia0, and DW PYTHIA), which are based on fits to data from several experiments spanning a wide range of energies is given. In addition, predictions of the PHOJET generator, which implements a different model of minimum bias events than PYTHIA, are discussed.

A full picture of high-energy hadron collisions typically combines perturbative QCD to explain parton interactions where it is applicable (high-$p_T$ scatterings), with an alternative phenomenological approach to describe soft processes, i.e. those with small momentum transfer $Q^2$. In soft interactions the strong coupling constant, $\alpha_s(Q^2)$, becomes too large for perturbation theory to be applied, and the parton interaction cross-section calculated by perturbative QCD models diverge as $p_T \to 0$.

Minimum bias final state is dominated by "soft" scattering processes. Soft processes also dominate particle production in the so-called UE, which is usually defined as all particles produced except those initiated by the hard process of interest in an event. Ideally, the measurement of minimum bias observables also constraints the UE models.

Examples of soft QCD models are the Dual Parton Model [20,21] and modified versions of QCD in which the divergencies arising from the running coupling constant are phenomenologically corrected to reproduce experimental observations. The MC generators PHOJET and PYTHIA implement these two different approaches.

Within PYTHIA a variety of different phenomenological models exist which all describe various experimental datasets if the parameters are tuned accordingly. Due to this flexibility a clear discrimination between models is often not possible. However, some observables do show a clear sensitivity to a specific model and will allow the results to be used in future tunings.

About 30% of the minimum bias events produced are of diffractive origin. Even though the fraction of diffractive events in the data sample is significantly reduced due to the event selection criteria, the remaining diffractive events usually have a small number of charged particles and build a significant fraction of events in this corner of the phase space. Hence they can not be neglected in the understanding of the minimum bias observables. PHOJET and PYTHIA implement different diffractive models.

To reduce the dependency on given diffractive model ATLAS decided not to corrected for these events as it is done in the so-called NSD cross-section which was done by many experiments before. Instead, the inclusive minimum bias cross-section for events with one track above $p_T$ of 500 MeV within the acceptance region of the central tracking system is measured. The measurement is fully corrected to hadron level. This means that the experimental results are directly comparable to the model predictions at the hadron level.

## 3.1 Samples generated for the minimum bias analysis

Some corrections in the data analysis chain can not be deduced from data alone, hence MC generators have to be used for it. In order to estimate possible model dependencies of these corrections, it is aimed to reach a maximal variation of phenomenological models and the variation of the predicted distributions while still being in agreement with previous measurements.

The soft QCD models describing soft hadronic interactions are implemented in various models in PYTHIA and PHOJET with an alternative ansatz. The free model parameters in the predictions of both generators have been tuned to data from LEP, HERA and Tevatron to cover the various aspects of particle production such as cross-sections, particle spectra and fragmentation.

In the following the models and tunes are described in detail with a special emphasis on possible observables in minimum bias collisions.

### 3.1.1 The PYTHIA model for soft hadron collisions

PYTHIA is a MC event generator program developed by the Lund group, frequently used for event generation in high-energy physics. A comprehensive description of the physics content of PYTHIA can be found in [8] and references therein.

PYTHIA contains a large variety of different scenarios and steerable parameters which have an impact on the description of the properties of minimum bias events. The core of PYTHIA's model for soft hadronic interactions is based on a phenomenological adaptation of QCD to describe the non-perturbative pp processes [8]. Variations on the main components of this model (e.g. multi-parton interaction, parton shower, treatment of beam remnants and colour reconnection) and their respective tuning to data have been developed and are available for different series of PYTHIA releases as can be seen in references [9, 10, 23].

The free parameters of the models inside PYTHIA have been tuned to many minimum bias and UE data samples spanning center-of-mass energies between 200 GeV and 1.96 TeV. It should be noted that there is no "optimal" tune which describes all existing data equally well. In particular, there seems to be a difference between the tunes optimised with minimum bias data only and those including also UE data. It is currently not clear, whether there is an inconsistency between the datasets or whether the models are not applicable equally well to both types of data.

To cover the full range of predictions within these uncertainties, we selected three different PYTHIA tunes for the minimum bias analysis in ATLAS, which vary both in the used models and in the datasets used for tuning.

The PYTHIA version used for the studies described in this thesis is 6.421. The following considerations should be taken into account when using PYTHIA to generate minimum bias events:

- Multiple parton interactions: PYTHIA allows the user to select different approaches for the handling of multi-parton interactions [8]. Distributions of charged particle densities, such as $\langle N_{ch} \rangle$ and $\mathrm{d}N_{ch}/\mathrm{d}\eta$ at $\eta = 0$, measured at different collider energies are fundamental to tune this component of the model. Prediction of the multiple parton interactions rate rise in LHC pp collisions at the design energy still remains as one of the dominant uncertainties of this model.

- The projectile matter distribution: Since each incoming hadron is a composite object consisting of many partons, there should exist the possibility of several parton pairs interacting when two hadrons collide. PYTHIA assumes that different pairwise interactions take place essentially independent of each other, and that therefore the number of interactions in an event is given by a Poissonian distribution. This is the strategy of the "simple" scenario [22].

  Considering that hadrons are not only composite but also extended objects, partonic interaction probabilities are likely to vary for each reaction. In this approach, called the "complex" scenario, the probability associated with each interacting parton depends on the assumed matter distribution inside the colliding hadrons. In the "complex" scenario an impact parameter ($b$) dependent approach is therefore introduced [22]. A small $b$ value corresponds to a large overlap between the two colliding hadrons, and hence an enhanced probability of multiple interactions. On the other hand, a large $b$ means a large probability that no pp interaction will take place in the event. This causes the so-called pedestal effect seen by UA1 [22] (the density of particles in the UE in jet events is larger than the density of particles of a typical minimum bias event) and also affects charge particle multiplicity distributions [23].

- Description of the beam remnant: the modelling of beam remnants is considerably different between the "old" and "new" scenarios. In both scenarios, all the scatterings that occur in an event must be correlated somehow by momentum and flavour conservation for the partons from each incoming hadron, as well as by various quantum-mechanical effects.

  In the old scenario, parton interactions are generated, Initial State Radiation (ISR) and Final State Radiation are estimated based on the virtuality scale and the resulting partons and the beam remnants pass through fragmentation to produce final-state hadrons. There is, therefore, a direct correlation between the number of multiple interactions and final-state hadrons [22].

  In the new scenario, in addition to the interleaved $p_T$-ordered showering model, the concept of colour reconnection is introduced [8].

- Color reconnection: The naive expectation from an uncorrelated system of strings decaying to hadrons would be that the mean transverse momentum $\langle p_T \rangle$ should be independent of $N_{ch}$. However, the general trend in collider data is that the tracks in high-multiplicity events are harder on average than in low-multiplicity ones. This can be explained by Color Reconnection.

Colour reconnections will typically occur between parton interactions and beam remnants, thus adding a new element in the model which is particularly sensitive to correlation between $p_T$ and multiplicity as well as the spread of the particles in rapidity. Appropriate tuning of the colour reconnection mechanism is proving to be essential to describe properties in minimum bias events, such as the $\langle p_T \rangle$ as a function of $N_{ch}$ spectrum, as shown in [10].

- Parton density functions (PDF): Each MC tune of parameters for minimum bias and UE properties is developed for a particular set of PDF. Thus, changing the PDF requires a re-tuning of the parameters.

For the MC production which started in 2009, ATLAS has decided to use the MRST LO* parton density functions [24] which provide a better description of processes using leading-order MC generators like PYTHIA. The ATLAS MC09 tune [9] is developed as an adaption of the MC08 tune which used CTEQ6L1 PDFs [25]. Both tunes use the $p_T$-ordered shower, double gaussian probability distribution of the matter distribution inside beam protons, and the strategy of color reconnection which minimises the total string length. In addition, the ISR and multiple parton interactions cut-off scales are separated from each other in MC09.

The model predictions are tuned to the published datasets of minimum bias and UE measurements from CDF Run I.

Samples generated using ATLAS MC09 PYTHIA tune are used to calculate detector acceptance efficiencies and to correct the data. In the Table 3.1 we show the generated cross-section for PYTHIA in the MC09 tune at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV.

| Process | Cross-section (mb) | |
|---|---|---|
| | 900 GeV | 7 TeV |
| ND (non-diffractive) | 34.4 | 48.5 |
| SD (single-diffractive) | 11.7 | 13.7 |
| DD (double-diffractive) | 6.4 | 9.17 |

Table 3.1: Generated cross-section for PYTHIA in the MC09 tune at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV . The statistical uncertainties on the numbers given are less than 0.1%.

## 3.2 The Monte Carlo models used to compare to the measurements in pp collisions

In this section we describe the various soft QCD and diffractive models as implemented in MC generators and different MC tunes on a level relevant for this analysis. Perugia0, DW and PYTHIA tunes are presented with a brief descriptions of the tunes, which are based on fits to data from several experiments spanning a wide range of energies. In addition, predictions of the PHOJET generator, which implements a different model of minimum bias events than PYTHIA, are discussed. It describes low-$p_T$ physics using the two-component Dual Parton Model.

### 3.2.1 Perugia0 tune

A set of tunes done by P. Skands using the $p_T$-ordered showers based on PYTHIA version 6.4 and CTEQ5L PDFs were presented at a workshop in Perugia [10]. These tunes have updated LEP fragmentation and flavour parameters and describe a large set of experimental data: Tevatron minimum bias data at $\sqrt{s}$=630, 1800, 1960 GeV and Drell-Yan[1)] data at $\sqrt{s}$=1800 and 1960 GeV, and SppS minimum bias data at $\sqrt{s}$ =200, 540 and 900 GeV.

The soft QCD part as modeled for the beam remnant and UE data are tuned using explicitly only minimum bias data from Tevatron and SppS. For these tunes, elastic and diffractive events are not simulated and less weight is given to low multiplicity events when tuning to data.

### 3.2.2 DW tune

The DW tune was derived using PYTHIA 6.2 to describe the CDF Run II underlying event and Drell-Yan data. It has 2.1 GeV of primordial $p_T$ [26]. This tune generates more high $p_T$ partons from both perturbative (shower) and non-perturbative (intrinsic) sources. Further details can be found in [26].

### 3.2.3 PHOJET

The PHOJET generator combines the ideas of the Dual Parton Model [20, 21] with perturbative QCD in order to describe hard and soft processes in a unified way [12]. The model employed by PHOJET is based on Regge theory, calculating the scattering amplitudes by taking the unitarization principle into account. PHOJET is formulated as a two component model, where the Dual Parton Model is used to describe the dominant soft processes and perturbative QCD is applied to generate hard interactions.

---

[1)] The DrellYan process occurs in high-energy hadron-hadron scattering. It takes place when a quark of one hadron and an antiquark of another hadron annihilate, creating a virtual photon or Z boson which then decays into a pair of oppositely-charged leptons. This process was first suggested by Sidney Drell and Tung-Mow Yan in 1970 to describe the production of lepton-antilepton pairs in high-energy hadron collisions.

Comparisons between the calculated results for cross-sections and the available data are used to determine the unknown model parameters (couplings, Pomeron intercepts and slope parameters), which are needed to generate multiparticle final states produced in inelastic interactions. It is therefore less flexible than PYTHIA regarding the tuning of its parameters.

PHOJET uses PYTHIA for the fragmentation of soft-chains (produced in Pomeron exchanges) or hard scattered partons based on the Lund model [8, 12]. This implies that any PHOJET tune is deeply linked with the particular version of PYTHIA. In this analysis, we use PHOJET version 1.12 and PYTHIA 6.421, a default in the current ATLAS software.

## 3.3 Summary

When bunches of protons collide in the LHC the most frequent interactions produce low transverse momentum particles. The physics of these interactions is not described by perturbative QCD. Since the underlying theory is lacking, an addition of phenomenological models is needed. These phenomenological models need have to be tuned to data and therefore rely on experimental inputs.

It should be noted that there is no optimal model or tune which describes all existing data equally well. Even though many minimum bias and UE data exist at energies between 200 GeV and 1.96 TeV to tune the models, many model uncertainties remain. These model uncertainties are partially caused by different datasets used for tuning. In particular, there seems to be a difference between the tunes optimised with minimum bias data only and those including UE data. It is currently not clear whether there is an inconsistency between the datasets or whether the models are not applicable equally well to both types of data.

To reduce the dependency on diffractive models ATLAS decided to measure the inclusive minimum bias cross-section for events with one track above $p_T > 500$ MeV within the acceptance region of the central tracking system and to present it corrected to hadron level. The non-diffractive, single-diffractive and double-diffractive contributions in the generated samples are mixed according to the generator cross-sections to fully describe the inelastic scattering.

# Chapter 4

# Event selection

This chapter describes selection criteria used by the ATLAS detector to measure the central pseudorapidity and transverse momentum distributions of charged particles produced in inelastic pp collisions at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV.

Events are first selected based on data quality information requiring the relevant detectors to be working with nominal performance. In addition a dedicated minimum bias trigger is required to have fired. Then basic event characteristics such as a reconstructed vertex and presence of at least one high-quality track are required. Tracks satisfying some quality requirements, that will be discussed in Chapter 5, form a sample of primary charged particles if they correspond to primary vertex.

The track and event selection used in the 7 TeV analysis is almost identical to that used in the 900 GeV analysis except for the addition of a veto on the multiple proton interactions per bunch crossing. The events are thus required

- to have all ID sub-systems at nominal conditions,
- to have passed the L1 MBTS_1 single counter trigger,
- to have a primary vertex,
- to not have a second primary interaction in the same bunch crossing (pileup veto) for 7 TeV analysis,
- to have at least one good track in the event. A good track is defined in the next Chapter dedicated to the track selection.

The pileup veto rejects events that have a second vertex with four or more tracks. It is described in more detail in Section 6.2.1.

## 4.1 Preselection of data

To better understand the underlying concept behind the software tools used for estimating the luminosity, one needs to consider the Luminosity Block (LB) concept of ATLAS. The smallest unit of ATLAS data is the LB. One LB contains roughly two minutes of data taking, but this can vary due to run conditions and other operational issues. The concept of treating the data in small chunks of time is not new [6]. It makes it possible to select data based on the quality and conditions of the beam and detector performance. The average luminosity in ATLAS is stored per LB. The trigger system of ATLAS provides information for the beginning and ending time of each LB. This information is recorded in databases which the users can access by time or LB number.

During the course of a given run only LBs with suitable run conditions are used. For minium bias measurements, the neccessary conditions are a fully functional L1 trigger and ID with all subdetectors operating at nominal voltage settings. The solenoid magnet is also required to be on.

#### The bunch crossing identifiers (BCID)

During early running of the LHC only few bunches were circulating in the LHC, both paired and unpaired, to be defined shortly after, crossings. The BCIDs of interest are those where:

- two beams are crossing at the centre of the detector. We call these collision or **paired BCIDs**. The BCID of paired bunches is known from the machine and only events with a BCID compatible with colliding bunches are selected;
- one beam is passing the centre of the detector but no beam in the other direction is present to interact with it. We call these single beam or **unpaired BCIDs**;
- there are additional BCIDs for some runs where there are collisions not far away from the centre of ATLAS but none in the centre of the detector. We call them near collision BCIDs.

This last sample contains a mixture of beam background and collisions remnants from collision happened not in ATLAS. We therefore remove those BCIDs from the analysis. The collision BCIDs are used for the analysis, the single beam ones are used to estimate the contribution of beam background events to our sample [18].

## 4.2 Trigger

The trigger requirements defined for this analysis are based on the MBTS already discussed in Section 2.2.1. The MBTS detector consisting of two discs situated on each side of the ATLAS detector, is used to define three trigger items: MBTS_1, MBTS_2 and MBTS_1_1. These items require a signal from either sides of the beam pickup, and respectively at least one MBTS hit, at least two MBTS hits and at least one MBTS hit on each side. A detailed description of these trigger items as well as a description of their performance can be found in [18] [27]. For this analysis, events are selected requiring the MBTS_1 trigger to have fired. This trigger choice has the highest efficiency [18].

## 4.3 Primary vertex

At least one primary vertex is required to be reconstructed by the vertex finders using at least three tracks. A vertex reconstruction efficiency is measured on data by counting events passing the trigger requirements that have a primary vertex reconstructed. This efficiency is found for events with at least one good[1)] track. The vertex reconstruction efficiency is also measured on MC events and found to be in agreement for events with at least one good track. The vertex reconstruction performance is described and studied in detail in [18] [27].

Occasionally, vertices can be reconstructed in background events corresponding to the interaction of the proton beam with residual gas in the vacuum of the beam pipe. Such vertices are uniformily distributed along the beam axis. To remove such vertices, a cut on the vertex $z$-position is applied.

The expected rate of the pileup events in the initial 7 TeV collisions is approximately $10^{-3}$, an order of magnitude larger than at 900 GeV. Therefore the algorithm attempts to reconstruct all primary vertices in each bunch crossing and relies on analysis-level requirement to select a single pp interaction.

## 4.4 The data taken at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV

The final counting of events and tracks with MBTS_1 trigger are presented in Table 4.1. After LB, BCID, trigger and offline selection[2)], we have a total of 326.201 events for a total of 1.863.622 tracks at 900 GeV. These numbers are identical to the published results [18].

The first run recorded by ATLAS at $\sqrt{s} = 7\ TeV$, run 152166 has 365.202 events and 3.741.256 tracks after applying all the requirements. This corresponds to more events and tracks than the full dataset used in the 900 GeV analysis. Therefore this analysis uses only this run. At 7 TeV we have almost the same number of selected events but a double of tracks as expected. These numbers are in agreement at 0.2% level with the public results [27]. The small discrepancy is due to small subset of data not found on the GRID [28].

| Data | Total events | Events after offline cuts | Tracks after offline cuts |
|---|---|---|---|
| 900 GeV | 948.350 | 326.201 | 1.863.622 |
| 7 TeV | 675.757 | 365.202 | 3.741.256 |

Table 4.1: Data sample used in this analysis after applied the offline cuts.

The number of events per LB after various stages of our selection process, so-called event flow, is shown in Figure 4.1. On the left we present one of the runs at 900 GeV, on the right the same plot for the run 152166 at 7 TeV, the only run used in the analysis.

The greened areas correspond to the good LBs. Each color corresponds to a different selection criteria applied. Pale yellow stands for all data recorded. Pale green shows the data with good LBs. Green shows the number of events that passed the collision BCID cut. Other gradation of green shows the number of events that passed the MBTS_1

1) A good track in this context is defined as a track satisfying the criteria described in this analysis where the requirements on transversal and longitudinal impact parameters, to be discussed in Section 5.1.4, extrapolated to the primary vertex are replaced by a cut on the transversal impact parameter with respect to the beam-spot requiring. See 6.2 Section dedicated to the vertex efficiency.

2) To be discussed in Chapter 5.

trigger, the vertex requirement and the track quality cuts. The tracks requirements are explained in details in the next Chapter. Dark greens shows the number of events that pass the impact parameters cuts. We call these cuts "primary". This is very close to the last dark green curve as the final selected sample requiring at least one good track (called "phase space"). The spikes and troughs in the LBs are due to detector operations that automatically change the LB number[3], such as changes due to the trigger pre-scales.

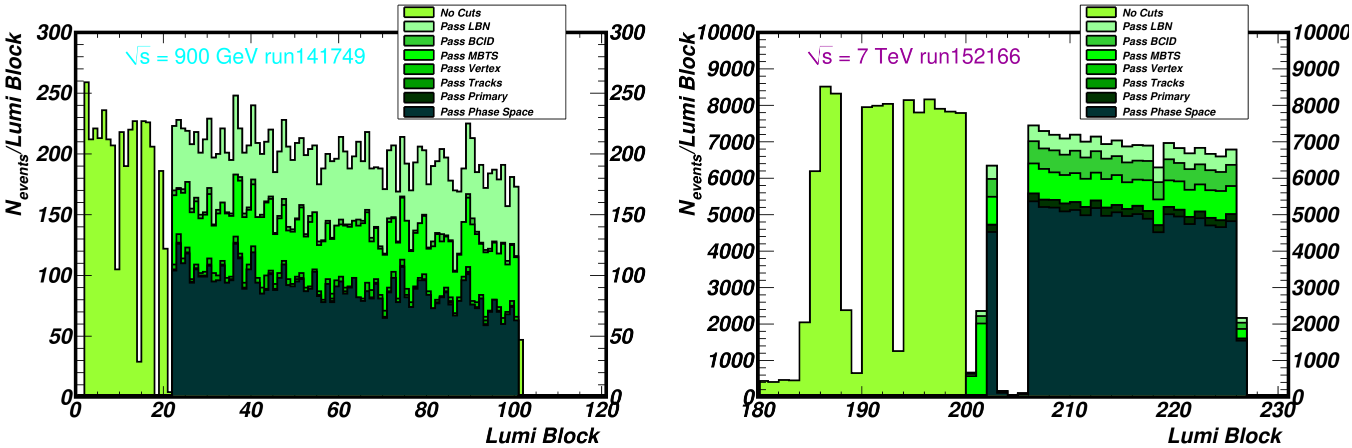



Figure 4.1: The number of events per LB after various stages of our selection process for one of the 900 GeV runs (on the left) and for the single 7 TeV run considered (on the right).

We simulated a million events at 900 GeV and at 7 TeV, with the cross-section weighted mixture of the non-diffractive, single and double diffractive processes. We use PYTHIA MC09 samples. The MC samples take in account the disable modules in the ID. Table 4.2 shows the fully operational part of ID during the data taking periods at 900 GeV and at 7 TeV. More modules were operational in the 2010 (7 TeV period) data taking than in 2009 (900 GeV period).

| Sub-detector | Fully operational at 900 GeV | Fully operational at 7 TeV |
|---|---|---|
| pixels | 96% | 98% |
| SCT (silicon strip) | 99.3% | 99.3% |
| TRT | 98.2% | 98.2% |

Table 4.2: Fully operational part of the ID during the data taking periods at 900 GeV and at 7 TeV.

## 4.5 The integrated luminosity

Luminosity at ATLAS is measured and monitored in several ways. There are absolute and relative luminosity measurements.

[3]The duration of an individual LB is not constant.

One type of absolute measurement uses the beam parameters to determine luminosity, e.g. van der Meer scans of the beam[4)]. Other absolute measurements make use of the optical theorem. The total rate of pp interactions is related to the rate of forward elastic scattering, Coulomb scattering, etc.; Roman pots will be used once the ALFA[5)] system [19] is installed.

The relative measurements are typically counting and/or coincidence experiments located at large angles (scintillators, high voltage current monitoring in calorimeters, LUCID [19]). Physics processes with high rate (e.g. W and Z production) can also be used to make luminosity measurements.

The precision of these measurements for the first run is strongly varying. The resulting luminosity values are uploaded to the conditions database and ultimately the physics users will be able to choose between the different methods.

Integrating luminosity over a fraction of data implies of getting the list of the corresponding LBs and, for each of them, to query the database for the information, then sum up the values using the simplified 'discretized integral formula':

$$L_{tot} = \sum_{i}^{LB} \triangle t_i^{trig} * L_i \tag{4.1}$$

where $L_{tot}$ is the total integrated luminosity, $i$ in the sum iterates through the selected list of LBs, $\triangle t_i$ is the time length of the $i$th LB of a given trigger, and finally $L_i$ is the average instantaneous luminosity of the $i$th LB from the database.

### 4.5.1 The integrated luminosity at 900 GeV

The integrated luminosity for the final event sample at 900 GeV, is estimated using a sample of events with energy deposits in both sides of the forward and end-cap calorimeters [29]. The luminosity presented here is that acquired by ATLAS and not the one delivered by the LHC.

The luminosity of pp collisions at 900 GeV is calculated as :

$$\int Ldt = \frac{N}{\sigma_{vis}} = \frac{N}{\epsilon_{ND}\sigma_{ND} + \epsilon_{SD}\sigma_{SD} + \epsilon_{DD}\sigma_{DD}} \tag{4.2}$$

where $\epsilon_{process}$ are the efficiencies and $\sigma_{process}$ the cross-sections for the individual inelastic processes contributing. These processes are the non-diffractive (ND), the single-diffractive (SD), the double-diffractive (DD). N is the number of observed collision events. This method relies on the understanding of these processes both in terms of their selection efficiency as well as their individual cross-sections. Both of them have potentially significant uncertainties.

The instantaneous luminosity is also determined per LB. It is simply given by the integrated luminosity in that LB divided by the length of the LB as stored in the database. The typical length of the LB during the 2009 data taking was two minutes.

---

[4)] In the Van der Meer scan one beam is swept stepwise across the other while measuring the collision rate as a function of beam displacement.

[5)] ALFA system consists of scintillating-fibre trackers located inside Roman Pots at a distance of 240 m from the ATLAS interaction point.

The total integrated luminosity for $\sqrt{s} = 900\ GeV$ is 19.2 $\mu b^{-1}$. The run with the largest integrated luminosity is 3.1 $\mu b^{-1}$, when the machine ran with sixteen colliding bunches. The luminosity during the stable beam period and that with the pixel and SCT bias voltage at nominal value is 12.2 $\mu b^{-1}$. Of this 9.1 $\mu b^{-1}$ were taken with the entire ID fully operational [29].

Figure 4.2 shows the instantaneous luminosity for the run 141994 at 900 GeV. The dashed line corresponds to the events selected with the MBTS_1_1 trigger, the light-shaded area to the events selected by the MBTS_1_1 trigger plus requiring paired bunches (BCID cut) and the dark area to the events selected by the MBTS_1_1 with the BCID and timing cuts applied[6)]. The red curve corresponds to an exponential fit to the luminosity lifetime, the fit parameters are given in the box. The bins with lower instantaneous luminosity correspond to short periods of time where the data acquisition system was suffering dead time.

In conclusion, at 900 GeV the luminosity of the data sample is approximately 9.1 $\mu b^{-1}$, while the maximum instantaneous luminosity is approximately $4.8 \times 10^{26} cm^{-2} s^{-1}$. The systematic uncertainty on reported luminosities is estimated to be about 20% dominated by the difference in modeling of diffractive components of the cross-section.

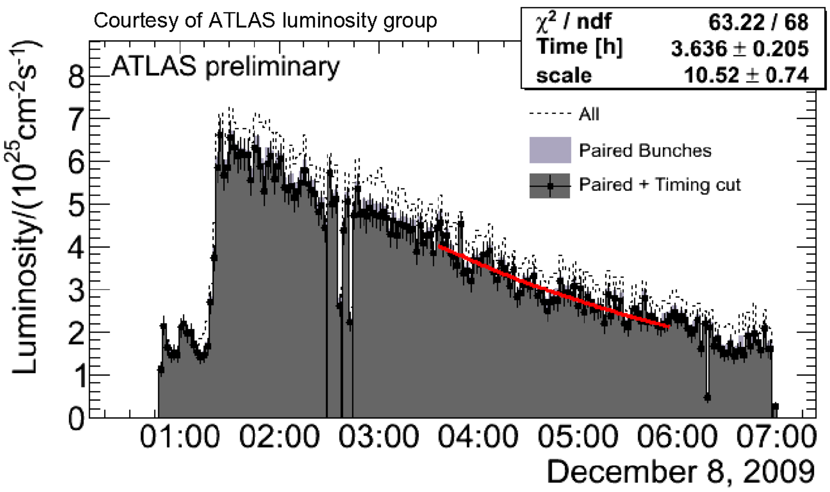



Figure 4.2: Instantaneous luminosity for the ATLAS run 141994 at 900 GeV.

### 4.5.2 The integrated luminosity at 7 TeV

At 7 TeV the luminosity was determined from counting rates measured by the luminosity detectors [29]. These detectors have been calibrated with the use of the Van-der-Meer beam separation method, where the two beams are scanned against each other in the horizontal and vertical planes to measure their overlap function.

6)By applying the timing difference cut between the A and C sides, to be defined in Section 7.2, of the MBTS scintillators for collision events we can reject beam-induced background events from un-paired proton bunch-crossings [18].

The run 152166 used for the analysis at 7 TeV corrisponds to 6.8 $\mu b^{-1}$ of data recorded by the ATLAS experiment. Figures 4.3 and 4.4 show the LHC beam energy, intensities and online luminosity as a function of LB number during run 152166.

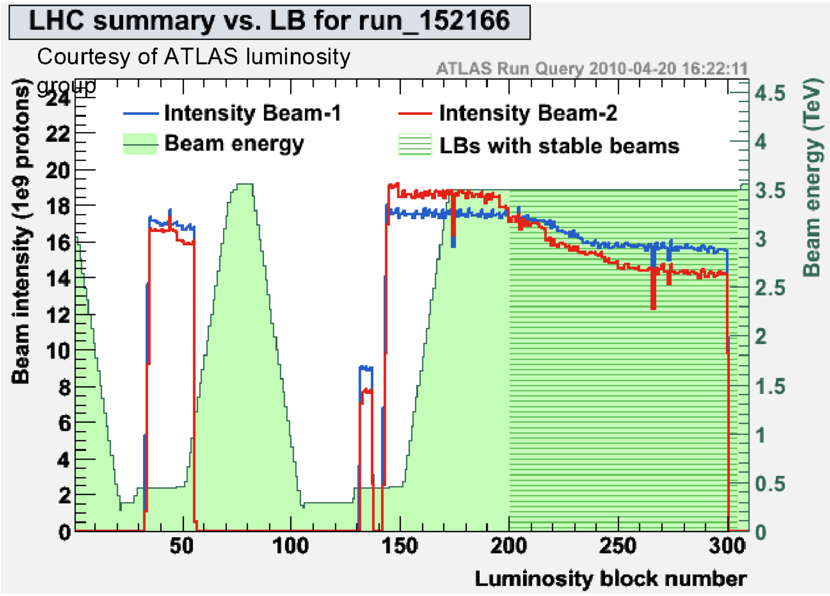



Figure 4.3: LHC beam energy and intensities as a function of LB number during run 152166 at $\sqrt{s} = 7\ TeV$.

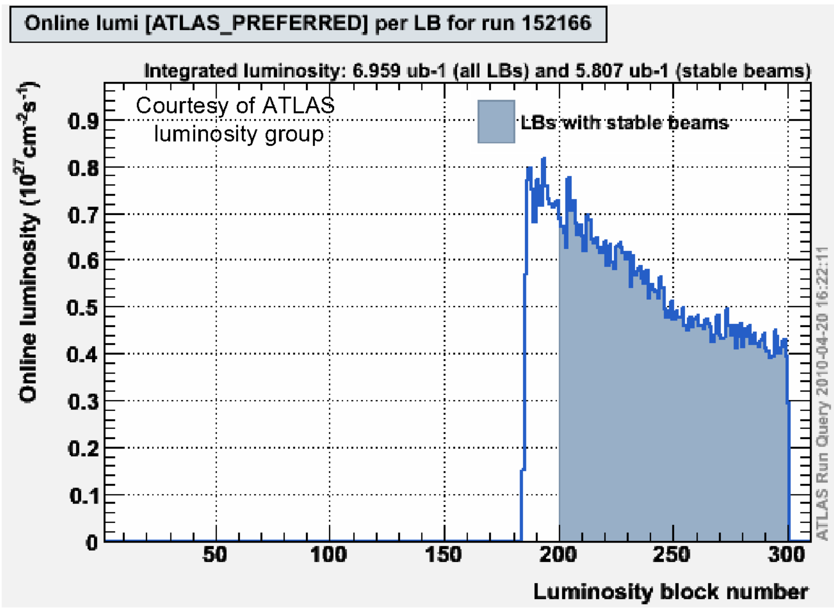



Figure 4.4: Online luminosity per LB for run 152166 at $\sqrt{s} = 7\ TeV$.

In conclusion, measurements of the LHC luminosity have been performed at $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$ using several detectors and methods.

The procedure used at 900 GeV had rough estimate of the systematic uncertainty. At 7 TeV we implement another method by cross-correlating the activities in the different forward detectors of ATLAS.

Absolute luminosity calibration obtained using efficiencies determined from MC are limited to about 20% by the systematic uncertainty associated with the modeling of the diffractive components of the pp cross-section, as well as by detector-specific systematic uncertainties that are typically of order of 5%.

Absolute luminosity calibration obtained from van der Meer beam scans at $\sqrt{s} = 7$ TeV has a systematic uncertainty of 11% which is dominated by the uncertainty in the bunch intensities [29]. The visible cross-sections obtained from the beam scans are lower than those predicted by PYTHIA MC09 (PHOJET) by 8% to 27% (26% to 35%) [29].

## 4.6 Summary

In this Chapter, we presented the methods to obtain the data samples which are used for the minimum bias analysis at 900 GeV and 7 TeV. We showed the LB and BCID selections, the selection cuts and the number of events at various stages of the event selection. No anomalous behaviour in the data is seen after the final event selection, demonstrating the consistency of the quality of the data recorded. Summarizing, events are selected according to the following criteria.

- Data quality: requiring the relevant sub-detectors, trigger systems and the solenoid magnet to be working with nominal performance. A selection of events based on the known ID's of the colliding bunches is also performed.
- Trigger: the MBTS_1 trigger is required to have fired.
- Primary vertex: a primary vertex reconstructed using at least three tracks is required to be present.
- Reconstructed track: at least one reconstructed track passing all track selection criteria is required to be present.

During this period of running, not all bunches in the machine were colliding at ATLAS. We must select, for analysis purposes, only those events from bunch crossing identifiers (BCIDs) corresponding to colliding bunches. For background studies we can use the BCIDs corresponding to single beams passing the centre of the ATLAS detector.

The ATLAS efficiency was about 90%. It is defined as the ratio between the data delivered with stable beam and the data recorded by ATLAS.

Accurate determination of luminosity is an essential ingredient of ATLAS physics program. In this Chapter we have shown the first results on luminosity determination, including an assessment of systematic uncertainties.

# Chapter 5

# Track selection

In the previous Chapter we discussed the event selection. Now we are going to present the selection of tracks. In accepted event at least one reconstructed track passing all track selection criteria was required to be present. This section details the requirements each track in the event should satisfy in order to be accepted as a high-quality track originating from the primary vertex.

Tracks that are accepted for the minimum bias analysis are required to fulfill the following criteria:

- Tracking algorithm: tracks are required to be reconstructed with the default inside-out tracking algorithm (further details in Section 6.3).
- Selection of good quality tracks consists of requiring at least one hit in the pixel detector and at least six hits in the SCT. This ensures the high quality of the selected tracks.
- Selection of primary tracks: $\mid d_0^{PV} \mid < 1.5$ mm, where $d_0^{PV}$ is the distance of closest approach to the primary vertex (PV) in the plane perpendicular to the beam direction. $\mid z_0^{PV} \sin\theta^{PV} \mid < 1.5$ mm, where $z_0^{PV}$ is the distance between the track and the primary vertex along the beam direction and $\theta$ is the polar angle with respect to the beam direction. This ensures that the selected tracks originate close enough to the primary vertex to be considered as primary tracks.
- Phase space definition consists of requiring $p_T > 500$ MeV and $|\eta| < 2.5$.

The track selection used for the 7 TeV analysis is identical to that used for the analysis at 900 GeV except for the addition of a veto on the multiple proton interactions per bunch crossing. The pileup veto rejects events that have a second vertex with four or more tracks; it is described in more detail in section 6.2.1.

## 5.1 Comparison between data and Monte Carlo

As the absolute tracking efficiency will be determined using the simulation, this section discusses the degree to which the simulation describes the data. Systematic uncertainties due to discrepancies between the data and the simulation will not be determined here, but rather from the impact on the tracking efficiency, which is discussed in Section 9.3.

The ID performance is illustrated in Figures 5.1 to 5.8 using selected tracks and their MC simulation. The distributions affected by overlapping pixel and SCT modules in the forward region and by inefficiency from a small number of disabled pixel modules in the central region are well modelled by the simulation.

### 5.1.1 Requirement on pixel hits

Requiring hits in the pixel detector is essential to have good resolution of the impact parameters which are used to identify primary tracks. Besides, this requirement helps removing secondary tracks originating from hadronic interactions by making sure that the selected tracks have hits in the detector layers close to the interaction point.

The fraction of disabled pixel modules plays an important role in the pixel detector performance. Modules that where disabled during part of the data taking are disabled in the reconstruction of all data runs and in the MC simulation in order to have an accurate and compatible description of the pixel detector performance. Figures 5.1 and 5.2 show distributions of the pixel hits and the average number of pixel hits per track as a function of $\eta$ and $\phi$. We present the same type of plots for 900 GeV and 7 TeV data samples.

The data are indicated with black dots, the ATLAS MC09 simulation in red histograms. We also report the ratio between data and simulation on the right plots.

The increase of hits on track at high $\eta$ shown in Figure 5.2 is due to the overlap between the modules in the barrel and the end cap, such that the possible number of hits per track increases. The remaining structure seen as a function of $\eta$ and $\phi$ is due to the disabled modules, but this is well described by the simulation. Good agreement between data and MC can be seen.

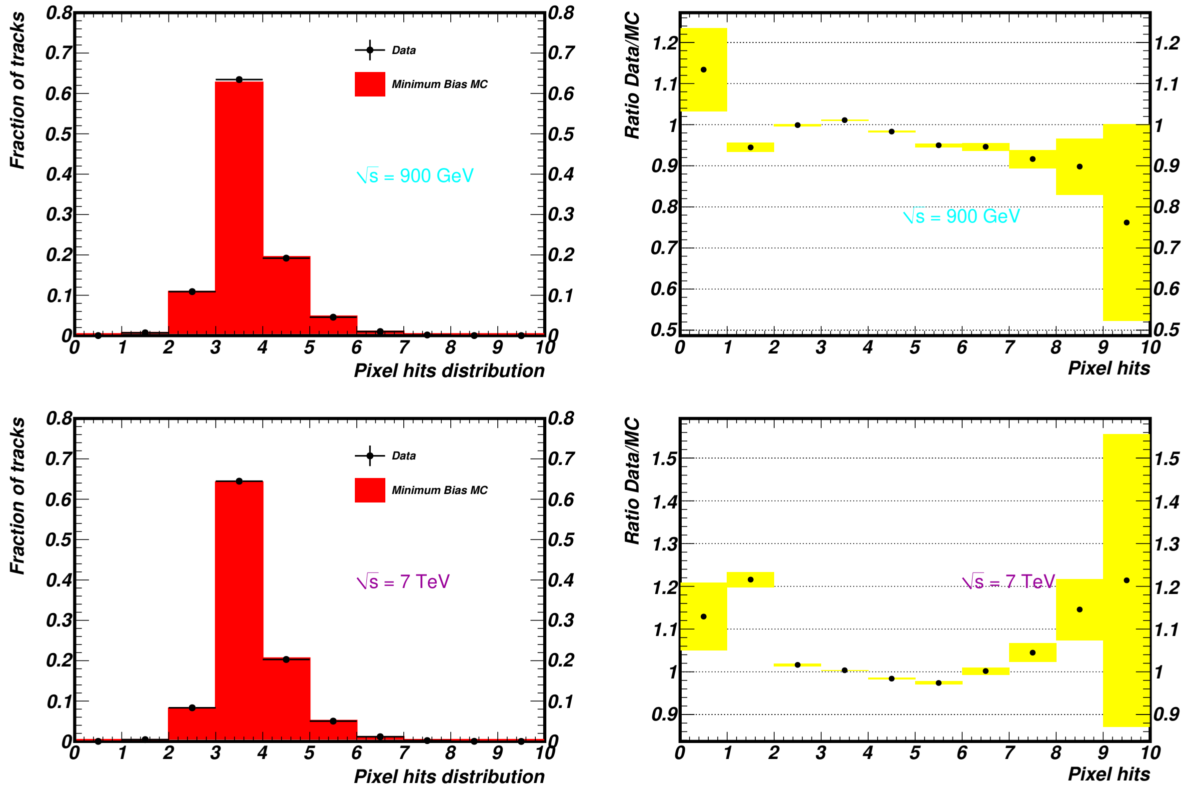



Figure 5.1: On the left top we give comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) for the number of pixel hits in units of probability at 900 GeV. On the right, ratios between data and MC simulation are shown. The yellow band stands for statistical uncertainty. On the bottom same plots for the 7 TeV data are presented. These distributions are done after applying the offline selection excluding the cut on the number of pixel hits.

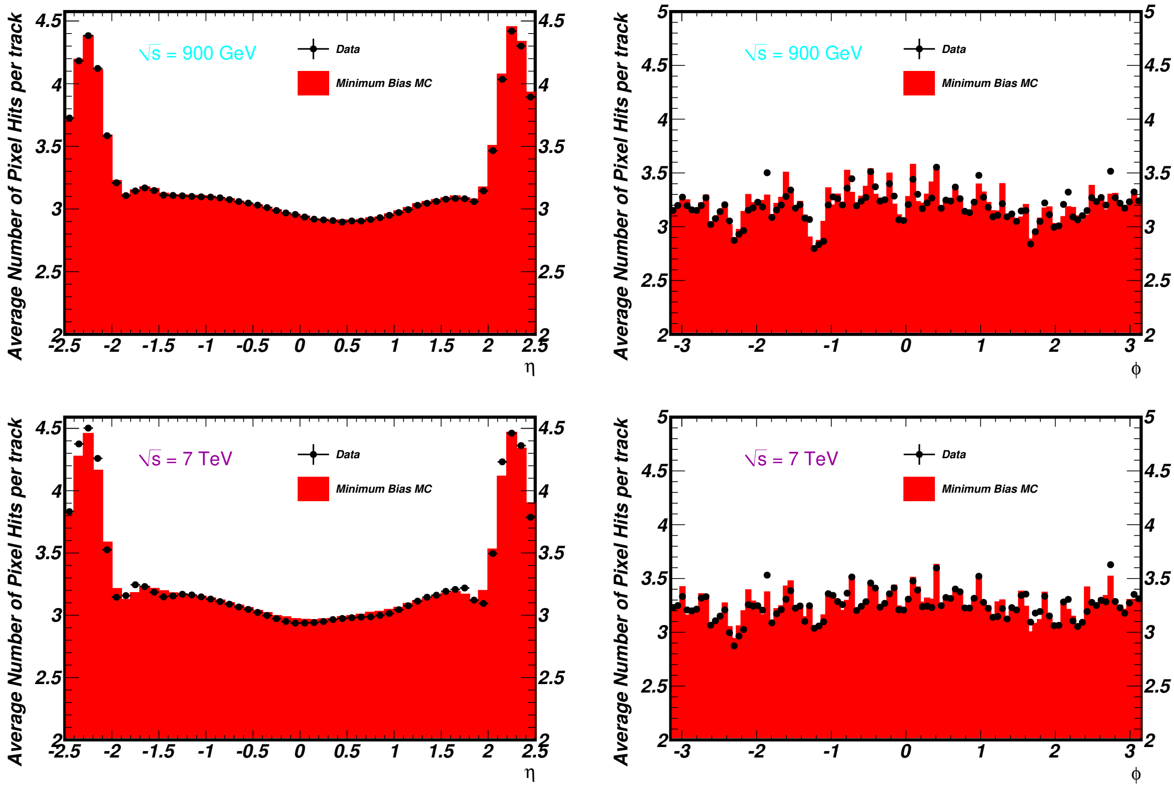



Figure 5.2: Comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) for the number of pixel hits as a function of $\eta$ in the left plot and as a function of $\phi$ in the right plot for 900 GeV data (top) and 7 TeV data (bottom).

### 5.1.2 Requirement on SCT hits

Requiring SCT hits on tracks insures enough hits at large radii to provide good resolution on the reconstructed track momentum. The SCT also serves as a major input to the ambiguity solving process, therefore reducing the number of fake tracks. For this analysis, we have chosen to accept only those tracks having at least six SCT hits. The performance of the SCT detector is found to be well reproduced by the MC simulation. Figures 5.3 and 5.4 show the multiplicity of SCT hits on track and the average number of SCT hits on track as a function of $\eta$ and $\phi$ for data and MC. A small disagreement of about 2% between data and MC can be observed in the forward region ($|\eta| > 2.2$).

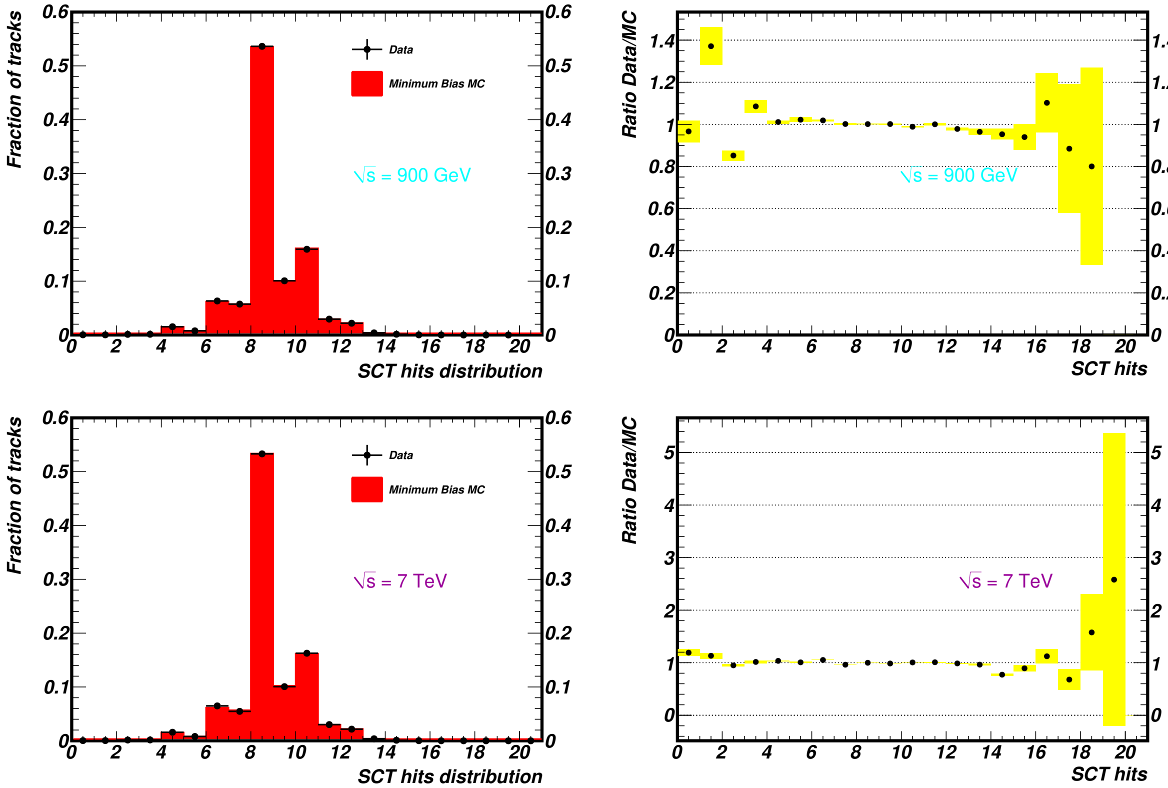



Figure 5.3: On the top left we present comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) for the number of SCT hits in units of probability at 900 GeV. On the right, ratios between data and MC simulation are shown. These distributions are given after applying the offline selection excluding the cut on number of SCT hits. Same plots are shown on the bottom for 7 TeV data.

Figure 5.5 shows the fraction of tracks passing different SCT hit requirements as a function of $p_T$ and $\eta$ for 900 GeV data. The efficiency of requiring at least six SCT hits is 97.5%. This efficiency is flat as a function of $p_T$. However, it shows dips in the efficiency as a function of $\eta$ corresponding to the transition between the SCT barrel and the end-caps where fewer measurements are possible.

The rejection of fake tracks and secondary tracks by above requirements is studied on MC events. No sizable effect is found on the rate of secondary tracks. The rate of fake

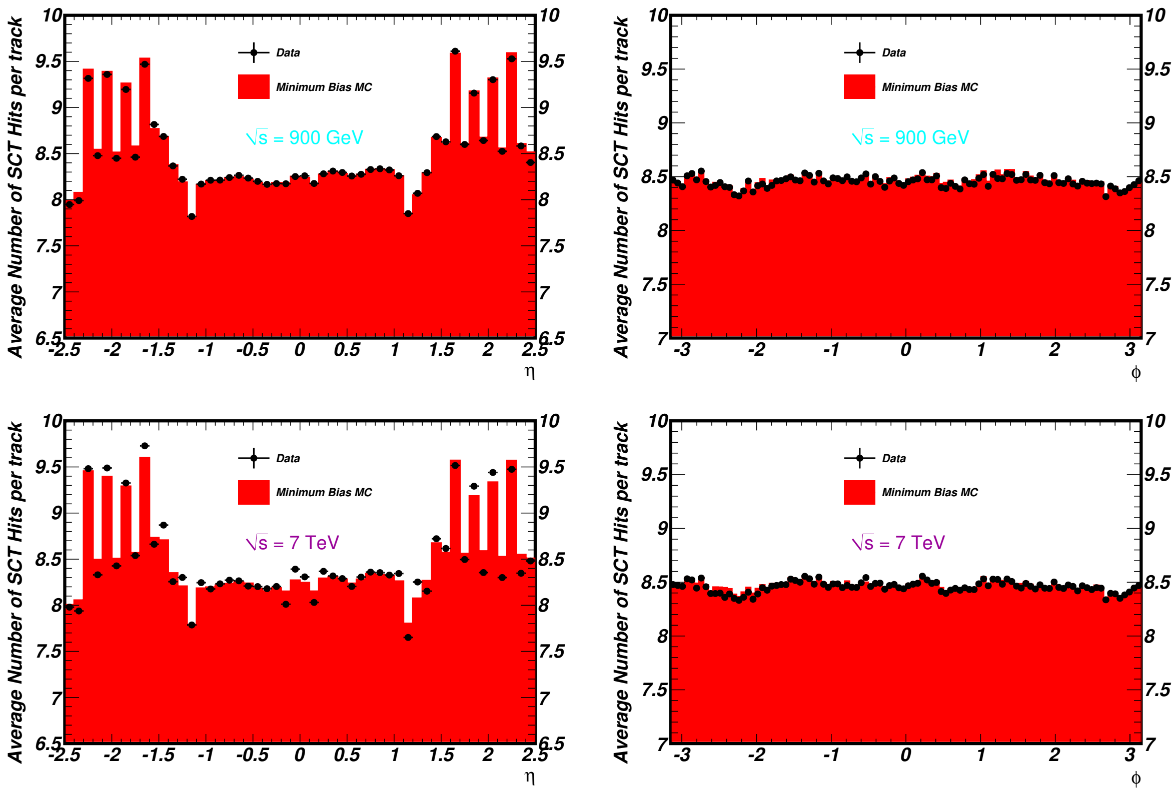



Figure 5.4: Comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) is shown for the number of SCT hits on track as a function of $\eta$ (left) and $\phi$ (right). As before, we present the 900 GeV data on the top and 7 TeV data on the bottom.

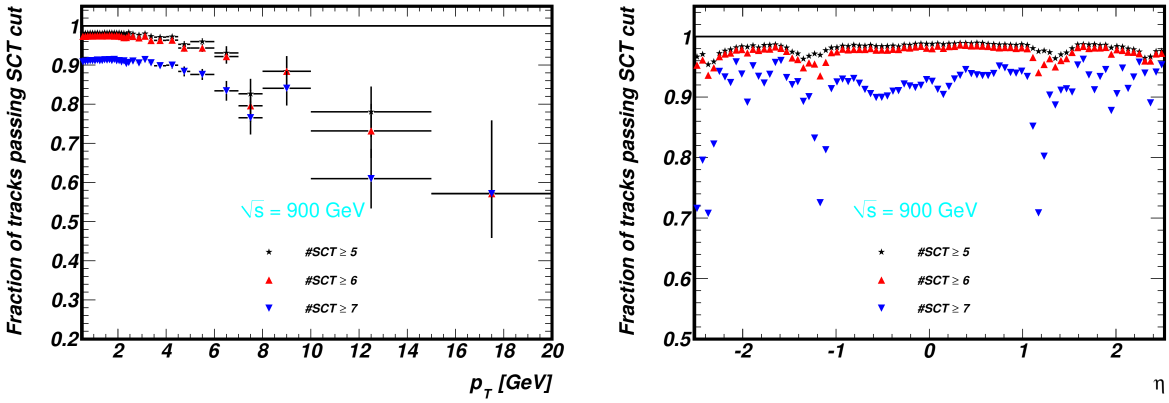



Figure 5.5: Fraction of tracks from the data passing all selection criteria and various SCT requirements as a function of $p_T$ (left) and $\eta$ (right), referenced to tracks selected without any requirement on the number of SCT hits.

tracks is found to be reduced by about 15% when applying the SCT hits requirement [18].

### 5.1.3 Requirement on TRT hits

The requirement of TRT extension was investigated. This requirement would further improve the resolution of the $p_T$ measurement by providing a series of measurement points in the transverse plane at large radius from the primary vertex. The effect of requiring TRT hits is studied in more details by comparing the properties of the tracks with and without TRT hits.

Figure 5.6 shows the number of TRT hits on track at 900 GeV and 7 TeV data. A disagreement between data and MC is observed, where on average, we see more TRT hits per track in the data. This disagreement was further investigated, and a bug in the digitization was found. The bug could not be corrected in time for this analysis, but it was corrected in the recent reprocessing of the MC.

About 10% of tracks don't have a TRT extension, that was shown in [18]. To allow a TRT extension, the inside-out tracking algorithm requires at least eight hits in the TRT which causes tracks that interact at the outer part of the SCT or at the inner part of the TRT to fail the TRT extension. These tracks can still be reconstructed with acceptable resolution using only silicon hits. The fraction of tracks with TRT extension is also well described in MC.

Apart from the overall 10% inefficiency when requiring a TRT extension, dips in the efficiency can be noticed at the TRT barrel to end-cap transition region and at $\eta = 0$ where extra passive material is present [18].

Preserving the 10% of the tracks without TRT extension is argued to be more important than the worsening of the quality of the collected tracks sample when no TRT hits requirement is applied. The TRT hits requirement is consequentally dropped from the selection criteria, although, we still use the TRT extension in the track fit whenever it is present to improve the momentum resolution. Dropping the TRT information completely even at the track fit level would have a rather large impact on the momentum resolution. On the other hand, the systematics on tracking efficiency are found to be negligible, while a small effect on momentum resolution is observed.

### 5.1.4 Impact parameters

Figures 5.7 and 5.8 compare the shape of the impact parameters distributions[1)] expressed at the primary vertex. The simulation describes the data well for $z_0$, however the $d_0$ distribution in the simulation is slightly narrower than that of the data. As the simulation sample is assuming a perfect alignment this small difference may be due to misalignment. As the impact parameter values are only used to suppress secondary tracks, a slightly different core resolution has negligible impact on the measurement (further details in Section 5.1.6).

From this comparison of basic tracking distributions, it appears that there is reasonable agreement between the data and simulation on the ID response to tracks. However, there are still a number of discrepancies that are yet to be understood, but have little impact under chosen track selection procedure.

Figure 5.9 (a) shows the fraction of tracks surviving different strength requirements on

[1)] Defined in Chapter 5.

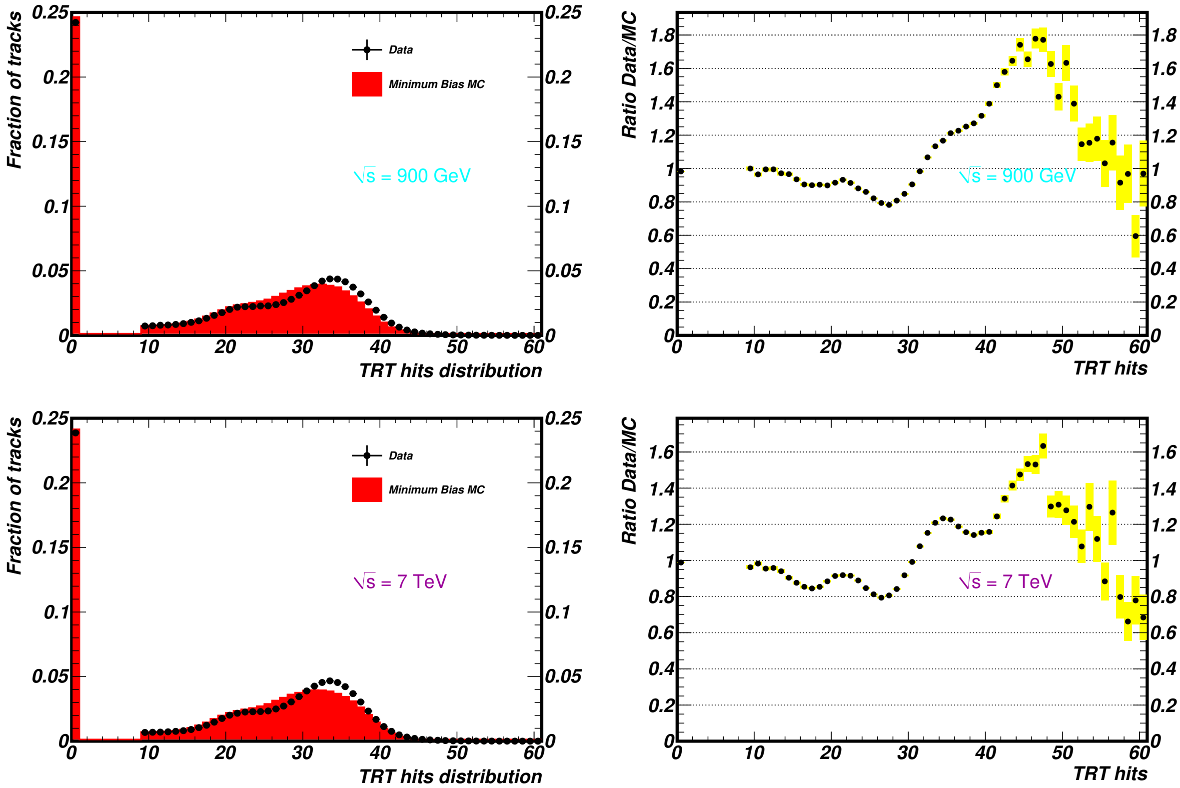



Figure 5.6: On the top left we show comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) for the number of TRT hits in units of probability at 900 GeV. On the right, ratios between data and MC samples are shown. These distributions are given after applying the offline selection excluding the cut on the number of TRT hits. Same plots are shown on the bottom for 7 TeV data.

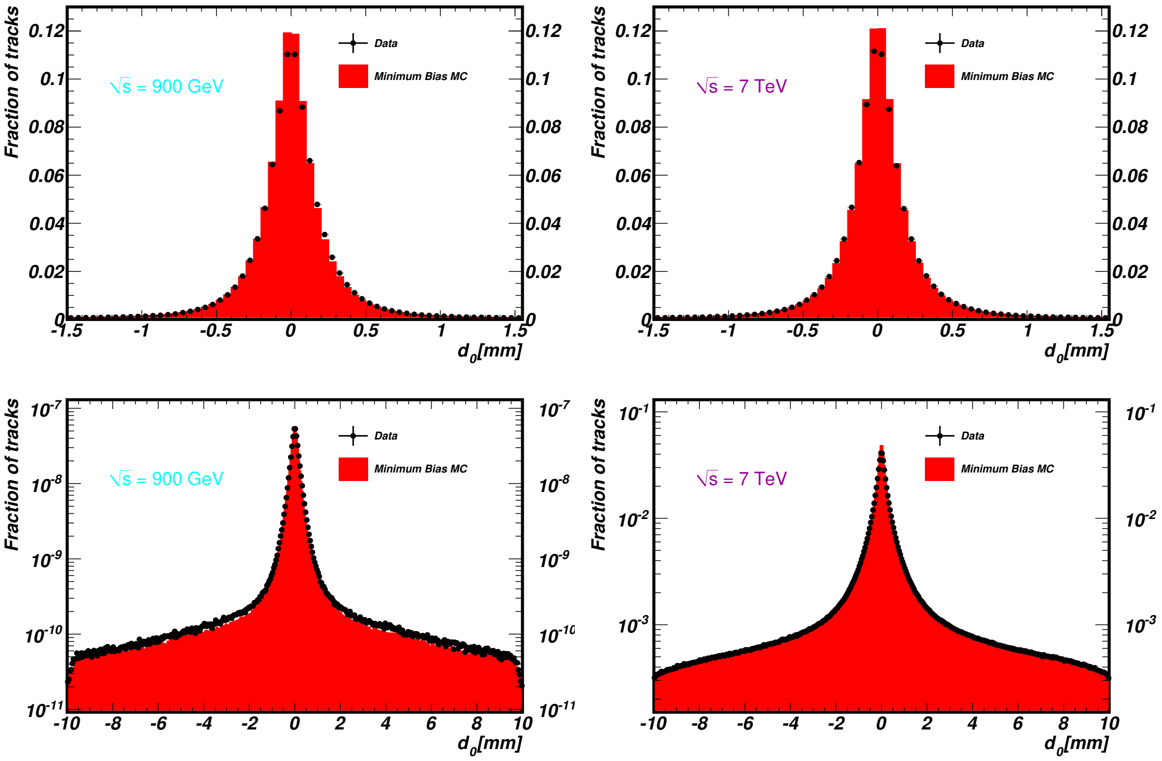



Figure 5.7: Comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) is shown for the $d_0$ distribution of the reconstructed tracks. In the top left plot we present the 900 GeV data, in the top right plot the 7 TeV track sample. The distributions are normalized to the total number of tracks in each sample. The bottom plots show the distributions in logarithmic scale and broader range of $d_0$.

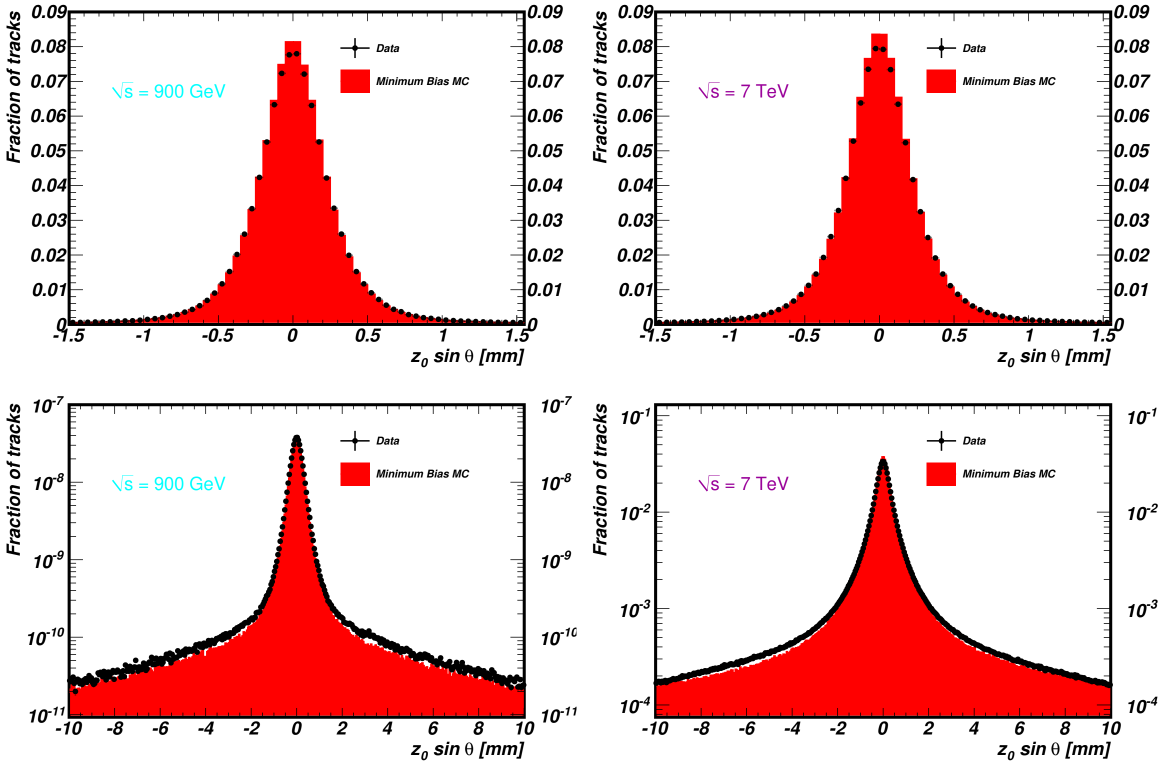
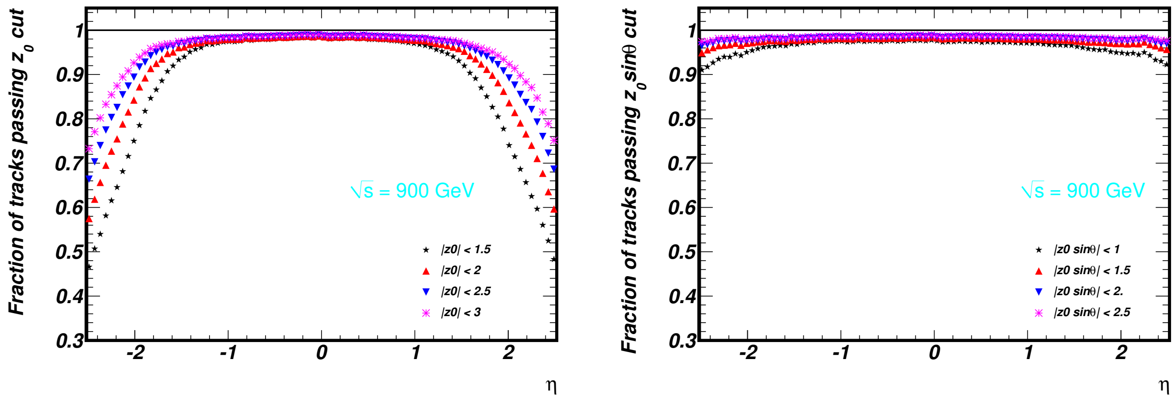



Figure 5.8: Comparison between data (dots) and minimum bias ATLAS MC09 simulation (histograms) is given for the $z_0$ distribution of the reconstructed tracks. In the top left plot we present the 900 GeV data, in the top right plot the 7 TeV track sample. The distributions are normalized to the total number of tracks in each sample. The bottom plots show the same distributions in logarithmic scale with a wider range up to 10 mm.

the longitudinal impact parameter as a function of $\eta$ as measured on 900 GeV data. Due to the fact that the impact parameter resolution gets degraded at high pseudorapidity, a significant inefficiency at large $\eta$ is observed. The ideal solution would be to apply an $\eta$ dependent cut which is tigher in the central $\eta$ regions and more loose at high $\eta$. Such a cut may also explicitly use the error on the impact parameter computed by the track fitting algorithms and cut on the significance of the impact parameter. This procedure needs more time to validate since it is very sensitive to the description of the material in the detector and the residual misallignements. The solution adopted for this analysis is to simply loosen the $d_0$ and $z_0 sin\theta$ requirements over the full $\eta$ range and to require $|d_0| < 1.5$ mm and $|z_0 sin\theta| < 1.5$ mm.

The effect of the impact parameter resolution at high $\eta$ is more pronounced for $z_0$ than for $d_0$ [18]. This is due to the fact that on top of the effect of material at high $\eta$, which equally affects $d_0$ and $z_0$, the $z_0$ resolution is also affected by the large lever arm that seperates the interaction point from the first measurement point. To correct for this effect, it was found necessary to define a cut on $z_0 sin\theta$ rather than on $z_0$. The result can be seen in figure 5.9 (b), which shows the fraction of tracks surviving several requirements on $z_0 sin\theta$ as measured on data.




(a) (b)

Figure 5.9: Fraction of tracks surviving different $z_0$ (a) and $z_0 sin\theta$ (b) requirements as a function of $\eta$ as measured on data.

Loosening the cut on the transverse and longitudinal impact parameters has an impact on the rate of secondary tracks in the selected sample. The rejection of secondary tracks is comparable for the $z_0$ and the $z_0 sin\theta$ requirements. The track selection efficiency is slightly better for the cut on $z_0 sin\theta$ especially at high $\eta$ [18], as clearly seen in Figure 5.9.

### 5.1.5 Comparison between 900 GeV and 7 TeV data

Figures 5.10 and 5.11 compare the number of hits on track between the 900 GeV and 7 TeV samples for each sub-detector in the ATLAS ID. The number of hits for the SCT agrees well between the datasets. However, differences are observed for the number of pixel hits. The increase in the number of pixel hits is because more modules were operational in the 2010 data taking than in 2009.

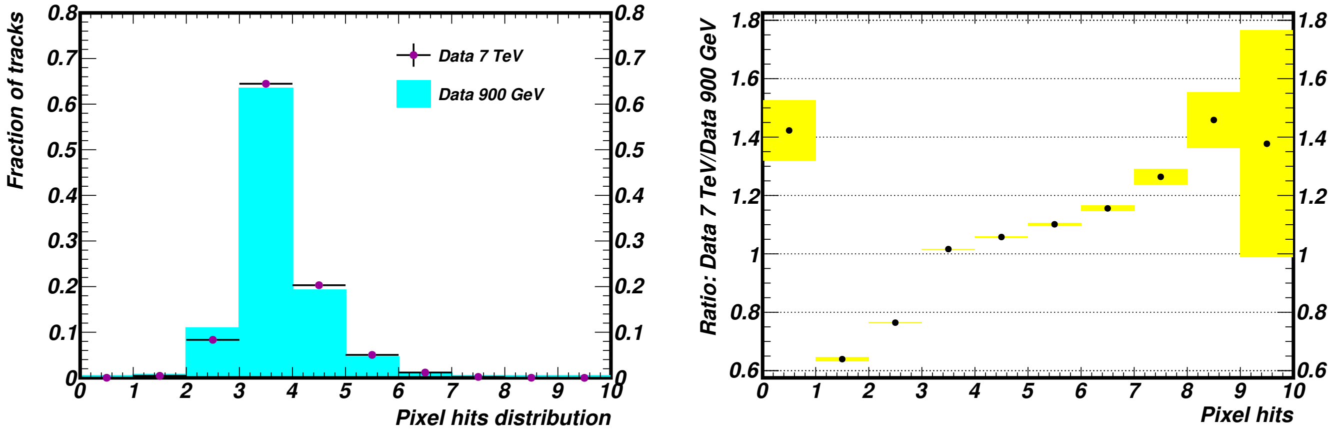



Figure 5.10: On the left we show the fraction of tracks as a function of the number of pixel hits in the 900 GeV and 7 TeV data samples. On the right we show the ratio between data at 7 TeV and 900 GeV as a function of the number of pixel hits. This distribution only depends on the detector conditions. The mean number of pixel hits has increased in 7 TeV data. The increase in the number of pixel hits is explained in the text.

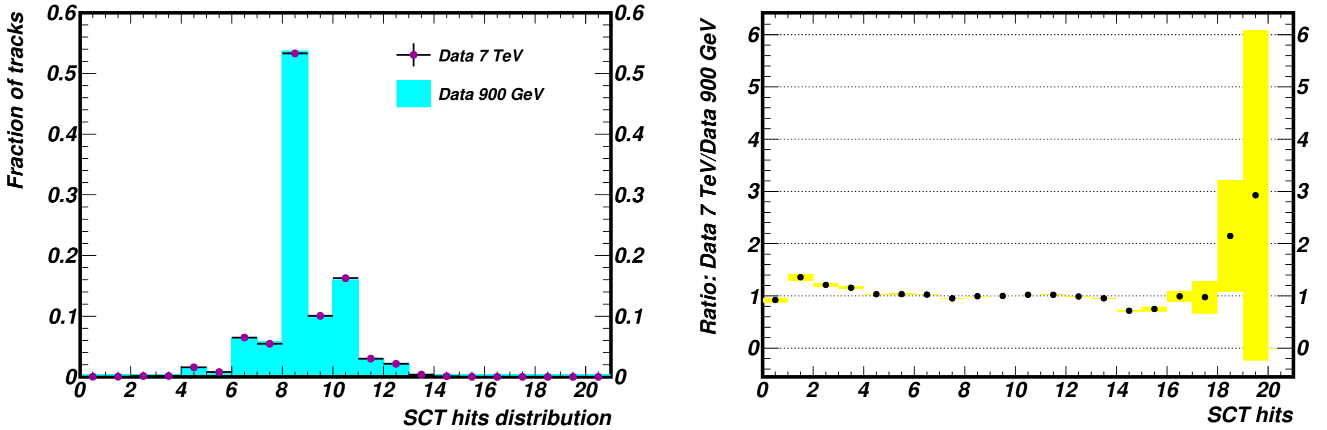



Figure 5.11: On the left the fraction of tracks as a function of the number of SCT hits in the 900 GeV and 7 TeV data samples is given. On the right we show the ratio between data at 7 TeV and 900 GeV as a function of the number of SCT hits. This distribution only depends on the detector conditions. The number of the SCT hits remains stable over few months that separate two datasets.

We also compare the raw distributions obtained from data and simulation for both energies. Tracks are selected as having measured transverse momentum ($p_T$), exceeding 500 MeV and an absolute value of pseudorapidity ($\eta$), below 2.5. The transverse momentum requirement corresponds to the threshold of the standard tracking algorithm, while the pseudorapidity requirement corresponds to the full acceptance of the ID. Figures 5.12, 5.13, 5.14 , 5.15, 5.16, 5.17 and 5.18 show the raw, detector level, distributions obtained from the data after event selection for the data samples at both energies. We present the distributions in different scales to amplify the tails.

The track distributions v*s.* $\eta$ and $p_T$ are shown in Figures 5.12, 5.13 and 5.14, respectively for data and MC at 900 GeV and 7 TeV centre-of-mass energies. The simulation sample is the cross-section weighted mixture of non-diffractive, single and double diffractive processes. Using the selected phase space and MBTS_1 trigger we eliminate a considerable part of diffractive events because they are usually very forward. The data are shown as black points, and the MC prediction, obtained using the PYTHIA MC09 tune, is shown as red shaded areas.

Table 5.1 lists the relative fraction of the components of the MC predictions at 900 GeV after applying the full event selection.

| Component at 900 GeV | Fraction remained after event selection | Generated cross-section |
|---|---|---|
| ND (non-diffractive) | 96% | 34.4 mb |
| SD (single-diffractive) | 2.7% | 11.8 mb |
| DD (double-diffractive) | 1.3% | 6.4 mb |

Table 5.1: Components of the total MC prediction as generated by PYTHIA in the MC09 tune, their relative fraction after event selection and the generated cross-section (used to scale the components relative to each other).

The lower and right plots show the ratio of data over MC as solid black dots and the statistical error as yellow shaded area. Both $\eta$ and $p_T$ distributions are in reasonable agreement between data and simulation and allow to estimate the efficiency as a function of both $\eta$ and $p_T$ (see Section 6.3).

Figures 5.15, 5.16, 5.17 and 5.18 show the number of selected tracks per event ($N_{tracks}$), where the tracks are those that pass all selection requirements, and the average transverse momentum $\langle p_T \rangle$ v*s.* $N_{tracks}$. In the bulk of the distributions data and MC agree within a few percent, while the tails show differences, up to a factor of two.

The distribution of $N_{tracks}$ agrees well except for the events with few tracks (Figure 5.16), where PYTHIA is expected to have difficulties to describe the data and where large corrections are applied.

### Choice of Binning

The size of the bins for all final observables is chosen such as to maximise the physics content of the plots while maintaining ease of readability. The smallest bin-size is chosen up to the limit where the statistics got close to the order of hundred events at which

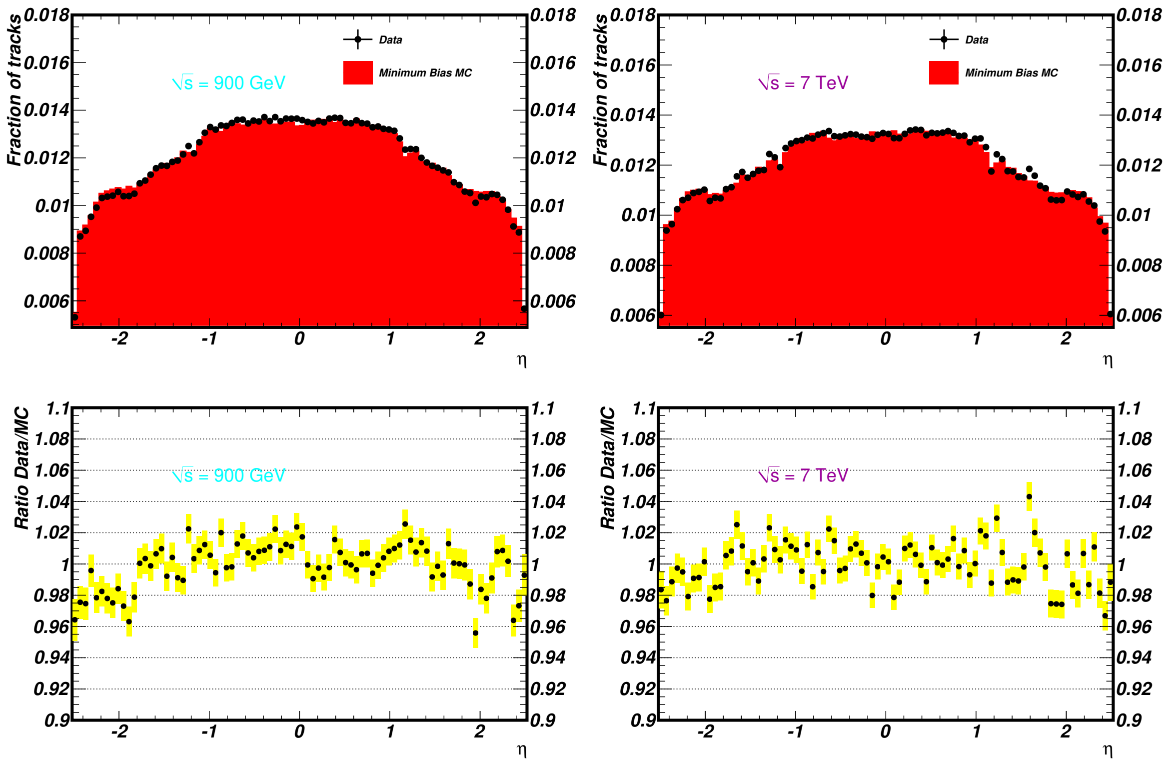



Figure 5.12: Uncorrected, detector level distributions of the track pseudorapidity are shown for 900 GeV data sample on the left and 7 TeV data sample on the right. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive components. The ratio of the data to the simulation is shown on the bottom plots.

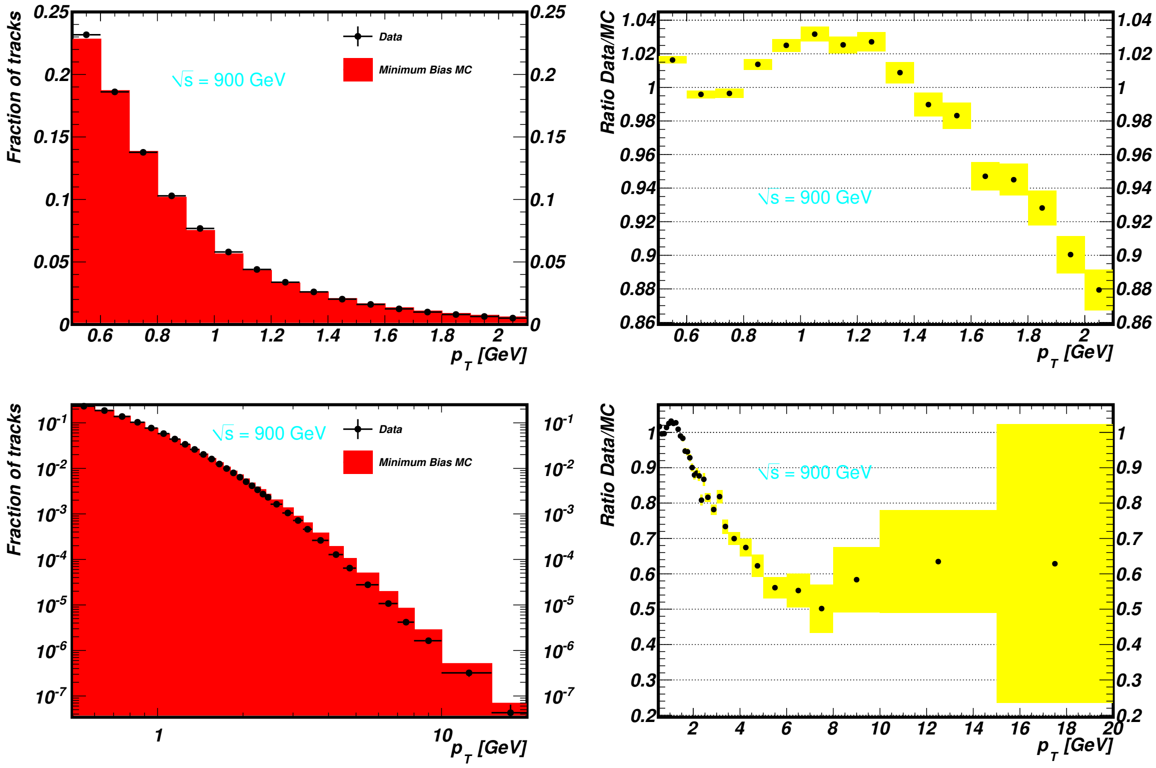



Figure 5.13: Uncorrected, detector level distributions of transverse momentum for 900 GeV data sample in linear scale are presented on the top and in logarithmic scale on the bottom. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive processes. The ratio of the data to the simulation is shown on the right plots.

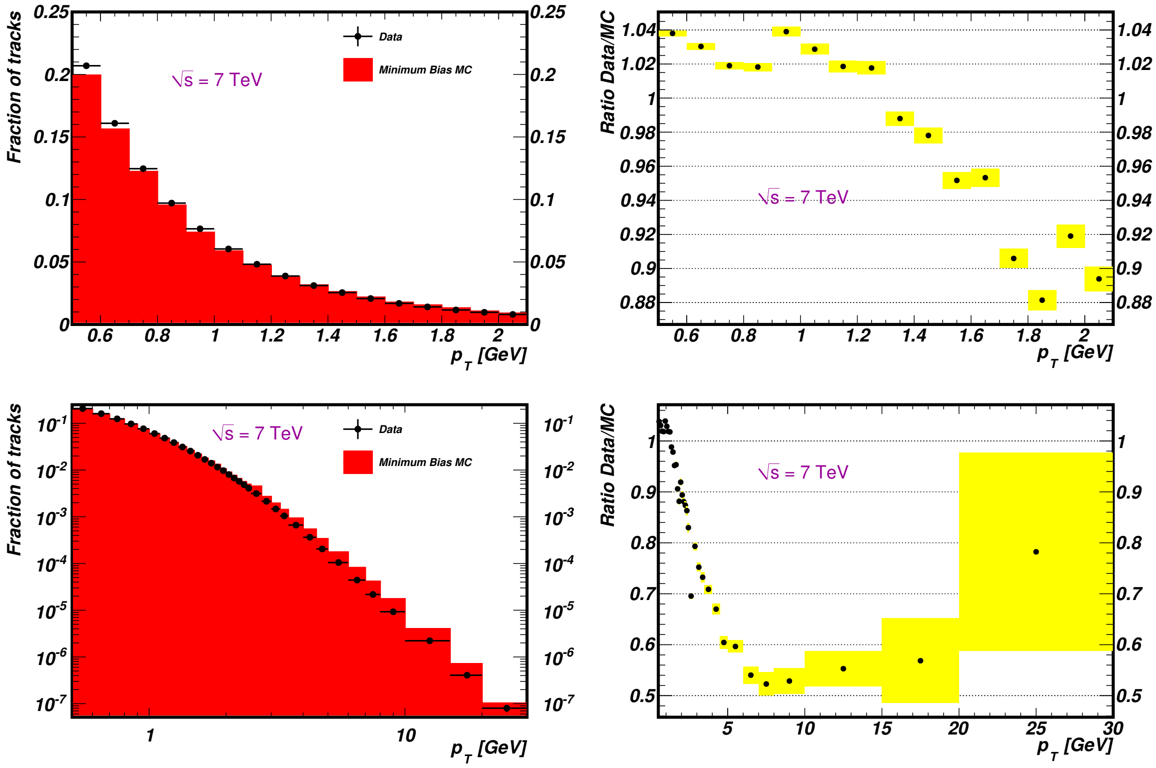



Figure 5.14: Uncorrected, detector level distributions of the track transverse momentum for 7 TeV data sample in linear scale on the top and logarithmic scale on the bottom. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive components. The ratio of the data to the simulation is shown on the right plots.

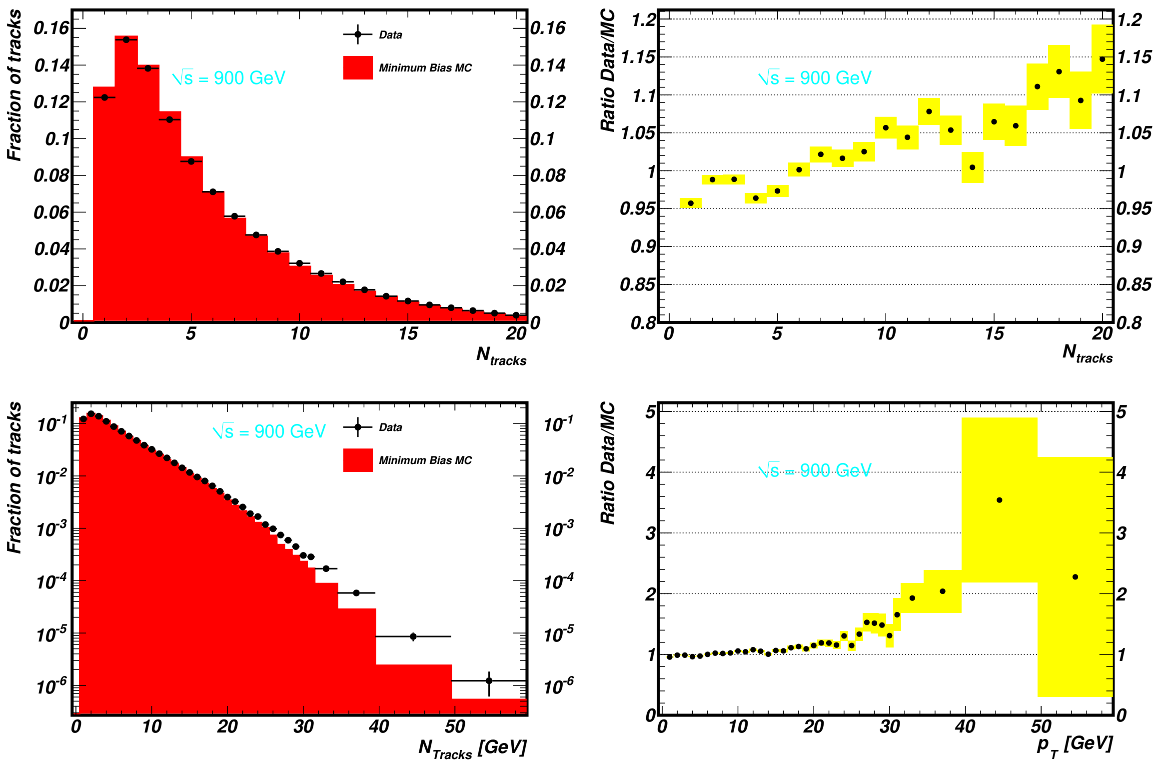



Figure 5.15: Uncorrected, detector level distributions of number of selected tracks per event for 900 GeV data sample are given in linear scale on the top and in logarithmic scale on the bottom. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive processes. The ratio of the data to the simulation is shown on the right plots.

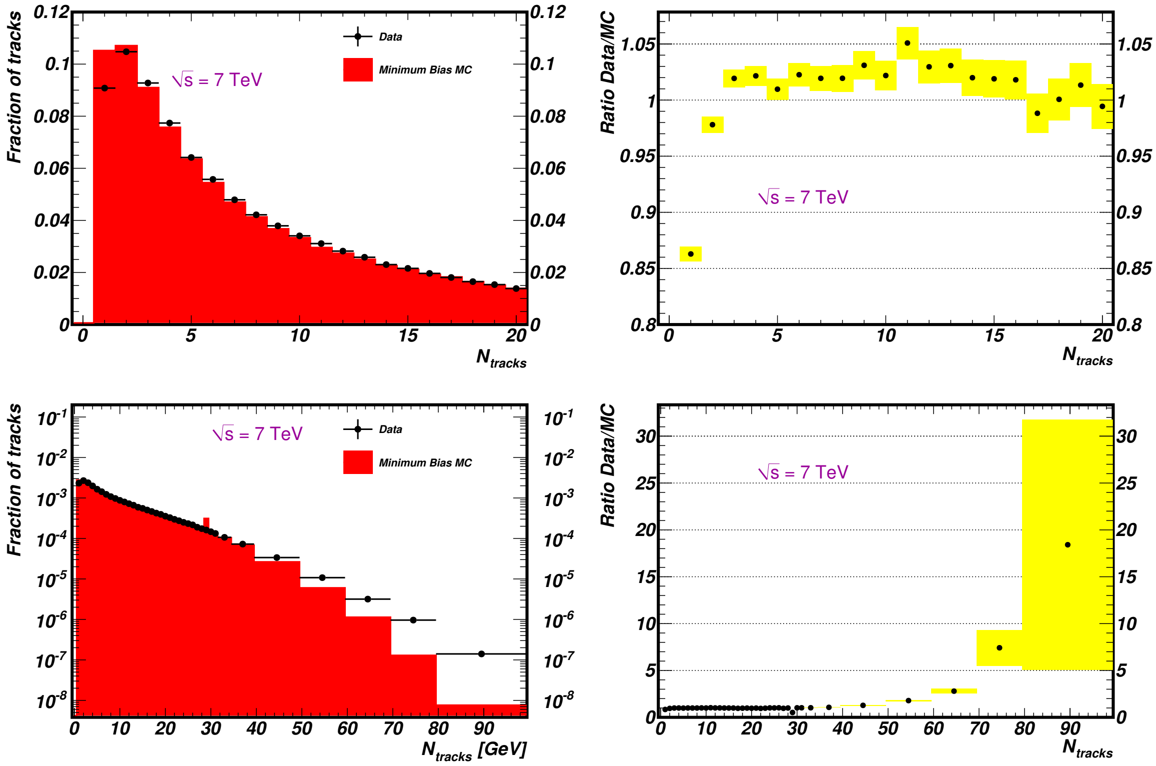



Figure 5.16: Uncorrected, detector level distributions of number of selected tracks per event for 7 TeV data sample in linear scale on the top and logarithmic scale on the bottom. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive components. The ratio of the data to the simulation is shown on the right plots.

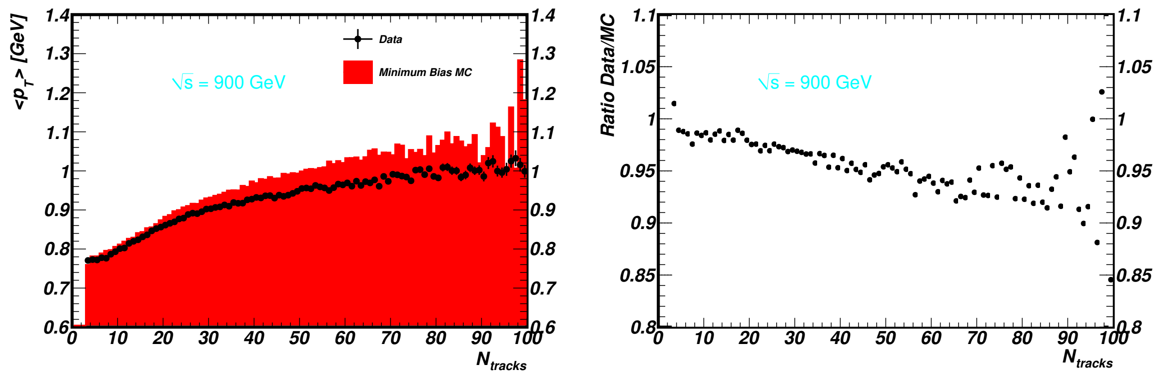



Figure 5.17: Uncorrected, detector level distributions of the average track transverse momentum is shown as a function of the number of tracks per event for 900 GeV data sample on the left and the ratio between data and MC samples is given on the right. The prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive processes.

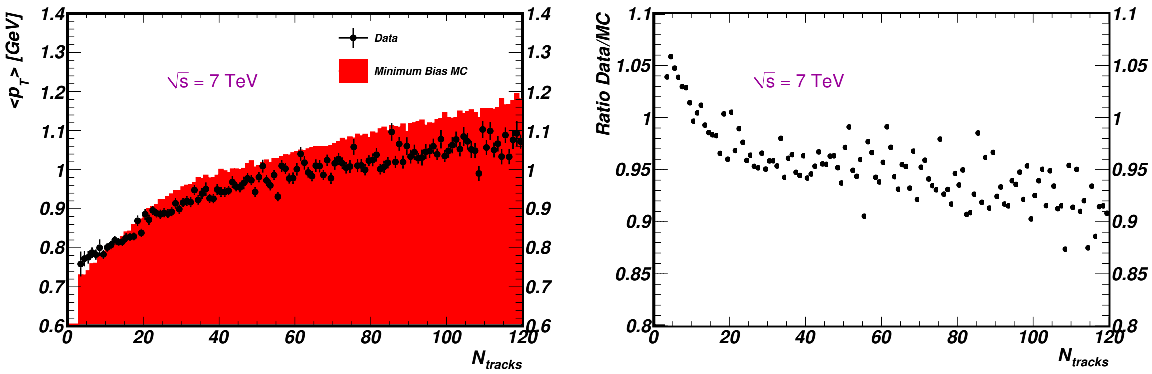



Figure 5.18: Uncorrected, detector level distributions of the average track transverse momentum is shown as a function of the number of tracks per event for 7 TeV data sample on the left and the ratio between data and MC samples is given on the right. The total prediction is calculated as the sum of the non-diffractive, single-diffractive and double-diffractive processes.

point the bins are widened usually by factors of two at a time, while trying to minimise the total number of bin-size changes, which are harder for the eye to grasp. The bins are extended until the data ran out of statistics. In some cases the last bin contains only four-six events. The bins are then fine tuned in distributions such as the $p_T$ spectrum. The statistical uncertainty is of similar size due to the difference between neighbouring bins. In the end the number of bin-size changes is small. Table 5.2 shows the final binning used for $p_T$, $n_{ch}$ and $\eta$.

| Variable | Bin Size | Range |
|---|---|---|
| $p_T$ [GeV] | 0.1 | 0.5 to 2.5 |
| | 0.25 | 2.75 to 3.5 |
| | 0.5 | 4.0 to 5.0 |
| | 1.0 | 6.0 to 10.0 |
| | then | 15.0, 20,0, 30.0 , 50,0 |
| $n_{ch}$ | 1 | 0.5 to 31.5 |
| | then | 34.5, 39.5, 49.5, 59.5, 69.5, 79.5, 99.5, 120.5 |
| $\eta$ | 0.1 | -2.5 to 2.5 |

Table 5.2: Binning used for the three different x-axes for 900 GeV and 7 TeV samples: $p_T$, $n_{ch}$ and $\eta$. The numbers correspond to the low bin edges for all but the last bin which also has the high edge shown as the last number.

### 5.1.6 Non primary tracks (fraction of secondary particles)

The distributions of charged particles in minimum bias events are determined for primary particles, i.e. those with a lifetime larger than $0.3 \times 10^{-10}$ s or decay products of particles smaller than those. I.e. pions, charged kaons or decay products of bottom and charmed hadrons are considered primary[2)]. However, the selected tracks used in the analysis include also non primary, i.e. secondary particles, which have to be corrected for. Secondary particles originate from different sources, e.g. from strange hadron decays, decays of pions and koans in flight, gamma conversions and hadronic interactions within the detector material.

Only primaries with $p_T$ >500 MeV and $|\eta| < 2.5$ are considered. More than 90% of the particles within this phase space are either pions or kaons.

This section describes how the fraction of secondaries in the selected tracks sample is estimated for data. The basic idea is to take the distribution of the impact parameter, $|d_0|$, i.e. the distance of closest approach to the event vertex in the plane orthogonal to the beam direction, at large values to obtain a rather pure sample of secondaries, assume the shape of the secondaries from simulation and adjust its yield on the data.

For these studies primaries are selected using the MC truth barcode[3)] for the truth particles. Primaries have the barcode between 0 and 200.000, secondaries are defined as having the barcode either equal to 0 or above 200.000. Fakes are defined as reconstructed tracks which do not match a truth track (further details in Section 7.3).

We are using the standard selection routines of the minimum bias analysis and therefore apply exactly the same selections as detailed in the previous sections, unless stated otherwise.

Figures 5.19 and 5.20 show the MC distributions for the distance of closest approach to the primary vertex for primaries secondaries and fakes at 900 GeV and 7 TeV, respectively. The MC samples are rescaled to total number of the data events that are represented by the black dots. Combined MC events are displayed in red. The primaries and secondaries are displayed by the blue and dark blue histograms respectively. The fakes are displayed by the green histogram. In the left plot you can see combined MC sample and separately secondaries and fakes, on the right we present the primaries, the secondaries and the fakes separately.

While the primaries are located almost exclusively in a narrow region of small $|d_0|$ with a width caused by the detector resolution, the secondaries produced in hadronic interactions and strange decays have a broader distribution and have to have an high purity in the region of $\mid d_0 \mid > 2$ mm.

We use the tails in the impact parameter distribution to estimate the fraction of secondary tracks in the minimum bias analysis for both centre-of-mass energies. The method is based on fitting MC templates to the data [18] [27]. The fit procedure has been also varied to study systematic dependences (see Section 9.6).

At 900 GeV the fraction of secondaries within the impact parameter cut is thus found to be $(2.20 \pm 0.05(stat.) \pm 0.11(sys.))$%. At 7 TeV the fraction of secondaries within

2) Pions produced from the decay of $\Lambda$ and $K_s$ are considered as secondary particles.

3) The barcode is a unique identifier used for book keeping in the generator and GEANT4. A barcode above 200.000 indicates the particle is handled by GEANT4. For the truth tracks a barcode of 0 indicates that the truth information was lost.

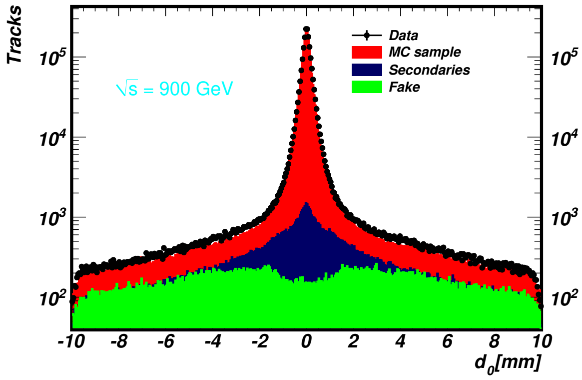



Figure 5.19: Comparison of $d_0$ distribution for the data (error bars) and MC for 900 GeV sample. The MC samples are rescaled to the number of the data events. Inclusive MC events are displayed in red. The primaries and secondaries are displayed by the blue and dark blue histograms respectively. The fakes are represented by the green histogram.

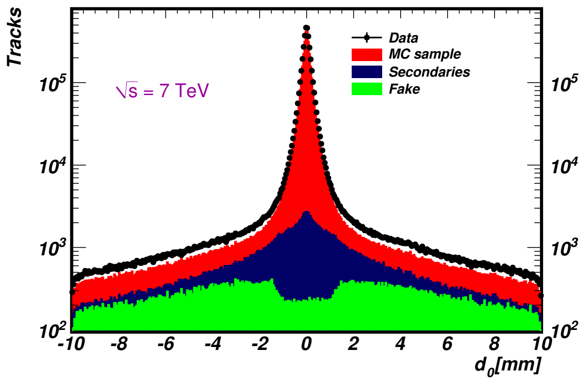



Figure 5.20: Comparison of $d_0$ distribution for the data (error bars) and MC for 7 TeV sample. The MC samples are rescaled to the number of the data events. Inclusive MC events are displayed in red. The primaries and secondaries are displayed by the blue and dark blue histograms respectively. The fakes are represented by the green histogram.

the impact parameter cut is thus found to be $(2.25 \pm 0.02(stat) \pm 0.11(syst))\%$. A small $\eta$ dependence of the fraction of the secondaries passing our selections is found, but it is within the systematic uncertainties and hence is not considered explicitly in the corrections in Chapter 8. The fakes result in negligible, below $10^{-3}$, contribution to the final data samples at both energies.

## 5.2 Summary

Data are selected by requiring one or more MBTS counter to be above threshold in coincidence with a collision bunch crossing identifier. Only those LBs in which the ID is fully operational are used.

The full acceptance of the ID is used and tracks are required to have $p_T > 500\ MeV$ and been reconstructed by the inside out track reconstruction algorithm [18] [27]. Events are required to have at least one reconstructed track and a reconstructed vertex containing at least three tracks at 900 GeV. At 7 TeV events with two reconstructed vertices with at least four tracks are rejected to reduce the contribution from pileup. Additional requirements are placed on the selected tracks to reject tracks produced by secondary interactions. The tracks are required to have at least one hit in the pixel detector and to have the transverse impact parameter, $|d_0^{PV}| < 1.5\ mm$ and the longitudinal impact parameter, $|z_0^{PV} \sin\theta| < 1.5\ mm$ pointing towards the primary vertex. To reduce the contribution from fake tracks, at least six hits in the SCT are required.

The number of events after each event selection criteria and the number of selected tracks in the final sample are shown in table 5.21 for three different types of trigger.

We observe an excellent agreement between data and MC. Thus we can use the MC to estimate the tracking efficiency.

| BasicChargedParticles: FirstCollisonAlg Output ========= | | | BasicChargedParticles: FirstCollisonAlg Output ========= | | |
|---|---|---|---|---|---|
| Cut Flow | Number of Events | N Quality Tracks | Cut Flow | Number of Events | N Quality Tracks |
| Total Events | 948350 | | Total Events | 675757 | |
| Events with Tracks | 948350 | 2184469 | Events with Tracks | 675757 | 19631936 |
| LBN | 530255 | 2112439 | LBN | 524739 | 18106991 |
| BCID+LBN | 461475 | 2112150 | BCID+LBN | 486759 | 18009551 |
| -------- | MBTS_1 | -------- | -------- | MBTS_1 | -------- |
| BCID+LBN+MBTS1 | 455593 | 2111926 | BCID+LBN+MBTS1 | 445697 | 18007910 |
| BCID+LBN+MBTS1+Vertex | 341256 | 2080524 | BCID+LBN+MBTS1+Vertex | 380442 | 17430114 |
| SCT+Pixel+author | 328745 | 2080524 | SCT+Pixel+author | 380432 | 9202744 |
| Primary track | 326571 | 1878501 | Primary track | 380314 | 7613877 |
| Phase Space | 326201 | 1863622 | Phase Space | 365202 | 3741256 |
| -------- | MBTS_2 | -------- | -------- | MBTS_2 | -------- |
| BCID+LBN+MBTS2 | 416738 | 2109061 | BCID+LBN+MBTS2 | 429964 | 18001913 |
| BCID+LBN+MBTS2+Vertex | 340220 | 2078700 | BCID+LBN+MBTS2+Vertex | 379943 | 17426954 |
| SCT+Pixel+author | 327884 | 2078700 | SCT+Pixel+author | 379933 | 9200672 |
| Primary track | 325718 | 1876759 | Primary track | 379815 | 7611999 |
| Phase Space | 325349 | 1861885 | Phase Space | 364801 | 3740525 |
| -------- | MBTS_1_1 | -------- | -------- | MBTS_1_1 | -------- |
| BCID+LBN+MBTS11 | 358579 | 2052506 | BCID+LBN+MBTS11 | 386675 | 17812028 |
| BCID+LBN+MBTS11+Vertex | 321140 | 2032437 | BCID+LBN+MBTS11+Vertex | 365045 | 17279905 |
| SCT+Pixel+author | 310727 | 2032437 | SCT+Pixel+author | 365039 | 9115652 |
| Primary track | 308769 | 1833689 | Primary track | 364932 | 7539552 |
| Phase Space | 308445 | 1819145 | Phase Space | 352411 | 3711374 |

Figure 5.21: Summary on the number of events and tracks passing each type of selection for 900 GeV on the left and for 7 TeV run on the right.

# Chapter 6

# Selection efficiency

The final distributions we are converging to are

$$\frac{1}{N_{\rm ev}} \cdot \frac{{\rm d}N_{\rm ch}}{{\rm d}\eta}, \quad \frac{1}{N_{\rm ev}} \cdot \frac{1}{2\pi p_{\rm T}} \cdot \frac{{\rm d}^2 N_{\rm ch}}{{\rm d}\eta {\rm d}p_{\rm T}}, \quad \frac{1}{N_{\rm ev}} \cdot \frac{{\rm d}N_{\rm ev}}{{\rm d}n_{\rm ch}} \quad \text{and} \quad \langle p_{\rm T}\rangle \ {\rm v}s.\ n_{\rm ch} \tag{6.1}$$

where $N_{\rm ev}$ is the number of events with at least one charged particle inside the selected kinematic range, $N_{ch}$ is the total number of charged particles, $n_{ch}$ is the number of charged particles in an event and $\langle p_T \rangle$ is the average $p_T$ for a given number of charged particles.

The data are corrected to obtain inclusive spectra for charged primary particles satisfying the event level requirement of at least one primary charged particle within $p_T > 500\ MeV$ and $|\eta| < 2.5$. These corrections include inefficiencies due to trigger selection, vertex and track reconstruction. They also account for effects due to the momentum scale and resolution, and for the residual background from secondary tracks.

Trigger and vertex reconstruction efficiencies are parameterised as a function of the number of tracks passing all of the track selection requirements except for the constraints with respect to the primary vertex. Instead, the transverse impact parameter with respect to the beam-spot[1)] is required to be less than 4 mm, which is the same requirement as the one used in the primary vertex reconstruction preselection. Since the beam dispertion along the $z$ axis is large, the longitudinal impact parameter at the beam-spot does not bring additional information regarding the rejection of secondary tracks. The multiplicity of these tracks in an event is denoted by $n_{Sel}^{BS}$.

Then in addition to the previous variables, definitions used throughout this analysis for the number of tracks or particles per event are

- $n_{Sel}$, the number of selected tracks per event, where the tracks are those that pass all selection cuts.
- $n_{Sel}^{BS}$, the number of tracks per event where the cuts with respect to the primary vertex $d_0^{PV}$ and $z_0^{PV}$ are not required but instead a cut on $d_0^{BS} < 4$ mm is applied, where BS means the track parameters are extrapolated to the beam-spot[2)].

[1)] The beam-spot position is computed by averaging vertex positions over a LB.

[2)] Note that the size of the luminous region (aka beam-spot) is not the same as the size of the LHC beams and the LHC bunch length, since it is given by the overlap integral of the two beams.

## 6.1 Trigger efficiency

Ideally the trigger efficiency would be found from a sample of events that satisfy the offline selection criteria and that are selected with the BPTX triggers. The only requirement of these triggers is the presence of beams passing through ATLAS, which means there is no requirement that an inelastic collision occured and no bias on the kinematics of the event. Unfortunately each of the two BPTX triggers is prescaled by a factor of about one thousand. The probability of an inelastic interaction in this dataset is $\sim 0.01\%$ and most events selected with this trigger contain no interactions.

An alternative method is to use a sample of events collected with the ID minimum bias trigger, called mbSpTrk and already discussed in Section 2.4. Then the efficiency of the single-arm MBTS trigger is determined from data using an independent trigger (mbSpTrk). It consists of a random trigger that requires only the event to be coincident with colliding bunches and some minimal activity in the pixel and SCT detectors.

Since mbSpTrk uses ID quantities the acceptance coverage overlaps with that for the offline cuts. Despite this overlap in acceptance, it is possible that events passing the offline event selection do not pass the mbSpTrk trigger. It is therefore necessary to investigate correlations in the mbSpTrk and MBTS triggers that could cause biases in the determination of the MBTS efficiency. This is discussed in [18] where only a small difference is observed. This is due to a fraction of events that pass the offline selection but not the mbSpTrk trigger. This effect is taken in to account during correction procedure in Chapter 8.

A sample of about 20K events that are selected on this trigger and that satisfy the offline selection requirements is used as the denominator of the efficiency calculation. The numerator contains the subset of these events that also fired the MBTS_1 triggers. The efficiency is therefore defined as:

$$\epsilon_{trig}(MBTS\_1) = \frac{MBTS\_1 \text{ AND } \texttt{offline} \text{ AND } \texttt{mbSpTrk}}{\texttt{offline} \text{ AND } \texttt{mbSpTrk}} \tag{6.2}$$

where `offline` denotes events that contain at least one reconstructed track passing the cuts listed in Chapter 4, and MBTS_1 (mbSpTrk) denotes that the MBTS_1 (mbSpTrk) trigger were fired in the event.

Equation 6.2 is used to determine the efficiency correction in each bin of the distributions of $n_{Sel}$, $p_t$ of the tracks and $\eta$ of the tracks, where $n_{Sel}$ is the number of selected tracks.

At the particle level, one needs to apply the corrections for the track distributions. The correction in bins of $p_T$ and $\eta$ is not an event efficiency as these are track level, not event level quantities. In this case the number of tracks in each bin satisfying the criteria in Equation 6.2 are counted and an efficiency correction is found. After correcting for the trigger efficiency $n_{Sel}$ distribution is corrected for the vertex and tracking efficiencies to obtain the fully corrected $n_{ch}$ distribution.

The trigger efficiency is shown in Figure 6.1 as a function of $n_{Sel}^{BS}$ for both datasets. The trigger efficiency is nearly 100% everywhere and the requirement of this trigger does not affect the $p_T$ and $\eta$ track distributions of the selected events. However, the trigger is slightly less efficient for low multiplicity events.

The trigger efficiency, $\epsilon_{trig}(n_{Sel}^{BS})$, is quoted as being dependent on the number of tracks per event meeting all analysis selection requirements except for the cuts on the impact parameters with respect to the primary vertex ($n_{Sel}^{BS}$). Instead, a cut at 4 mm is applied on the $d_0$ with respect to the beam-spot.

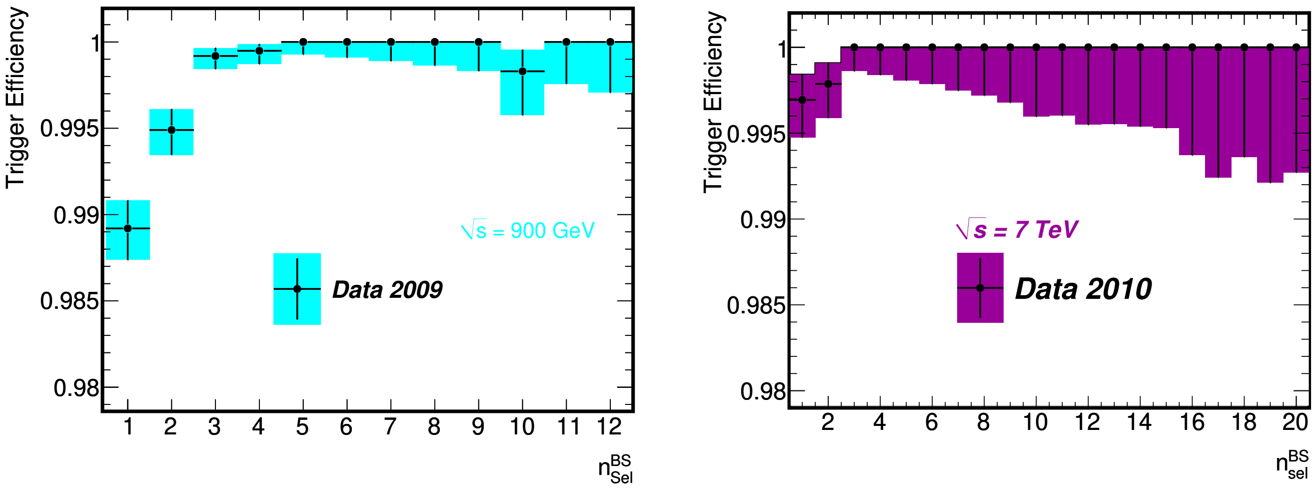



Figure 6.1: Trigger effiencies as a function of the number of tracks passing all of the track selection requirements in respect to the beam-spot for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot.

Several changes have been made to the MBTS trigger between 2009 and 2010 data-taking [27]. The photomultiplier high voltage of the MBTS counters has been increased from their nominal value of 700 V to 850 V, which moves the location of the MIP peak from 0.15 pC to 0.6 pC. The trigger threshold is raised from 30 mV to 50 mV so that a single MIP trigger signal is accepted with higher efficiency. The trigger item used for the selection of the data MBTS_1 has been redefined for the 7 TeV data sample. It requires a coincidence of a filled bunch crossing with one or more MBTS counters above threshold, however this has no impact on the analysis. For these reasons we observe higher efficiency at 7 TeV (Figure 6.1). A separate trigger, MBTS_1_UNPAIRED, has been introduced which is used to study the beam background.

## 6.2 Vertex reconstruction efficiency

The reconstruction of primary vertices plays a central role in studies of particle multiplicities in minimum bias $pp$ collisions. One of the most important selection criteria for the charged tracks produced in the primary interaction are their transverse and longitudinal impact parameters, $d_0$ and $z_0$, respectively. The latter quantities (or their corresponding significances) are usually calculated with respect to the reconstructed primary vertex and allow the separation of primary and secondary tracks. The various effects influencing the efficiency of the primary vertex reconstruction thus introduce important systematic errors to the final measurement of the charged particle multiplicity.

The vertex reconstruction efficiency is determined from the data, by taking the ratio of triggered events with a reconstructed vertex to the total number of triggered events. It is shown in Figure 6.2 for both data samples as a function of $n_{Sel}^{BS}$, the same track definition used to parametrize the trigger efficiency.

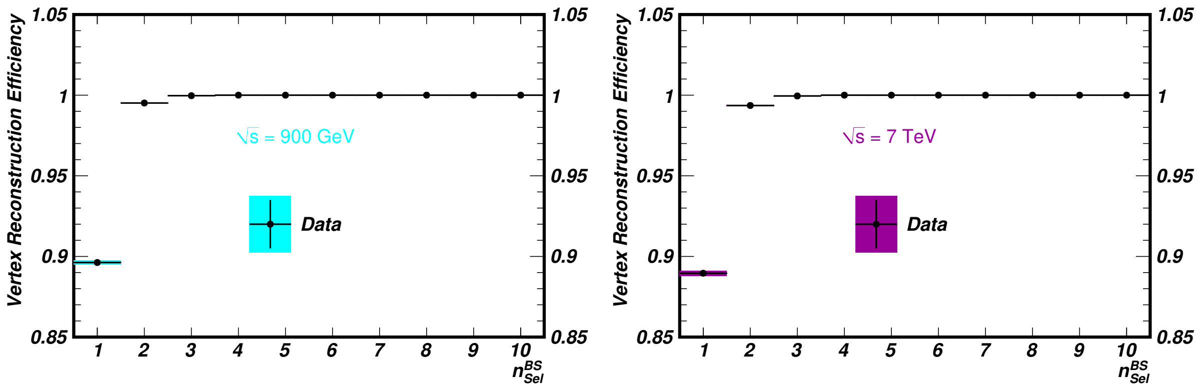



Figure 6.2: Vertex reconstruction efficiencies as a function of the number of tracks passing all of the track selection requirements in respect to the beam-spot for $\sqrt{s} = 0.9\ TeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. In both cases the same algorithm was used.

The strategy for the reconstruction of primary vertices in the 7 TeV run differs significantly from the approach used for 900 GeV published analysis [18]. However both data samples used an iterative adaptive finder algorithm that has the ability to reconstruct multiple primary vertices. In addition, the constraint that the reconstructed vertices must be consistent with the beam-spot is used.

Compared with the published 900 GeV data the only change to the selection of the tracks that are considered by the reconstruction algorithm is the lowering of the $p_T$ threshold from 150 to 100 MeV. The remaining preselection cuts are identical to the 900 GeV analysis. The beam-spot information is used both in track preselection and to constrain the fit during the vertex reconstruction process. Because of the additional constraint of the beam-spot, the requirement of there being at least three tracks in the vertex was removed. The algorithm itself has a limitation on a minimum of two tracks per vertex.

With the new vertex reconstruction algorithm and the loosening of the track requirement in the vertex from three to two, the efficiency is higher for low multiplicities. For events containing fewer than three selected tracks, the efficiency is found to depend on the pseudorapidity. Therefore the vertex reconstruction efficiency is corrected in bins of $\eta$ for such events, but not for higher multiplicity events. No dependence on $p_T$ is observed.

In addition, the expected rate of the pileup events in the initial 7 TeV collisions is approximately $10^{-3}$, an order of magnitude larger than at $\sqrt{s} = 900\ GeV$. Therefore the algorithm attempts to reconstruct all primary vertices in each bunch crossing and relies on analysis level requirement to select a single pp interaction.

### 6.2.1 Pileup effect

The rate of multiple interactions in the same bunch crossing is expected to be of the order of $10^{-3}$ for 7 TeV run. Therefore the impact of pileup is investigated.

Figure 6.3 shows the ATLAS event display of pileup event, where we observe two primary interactions.

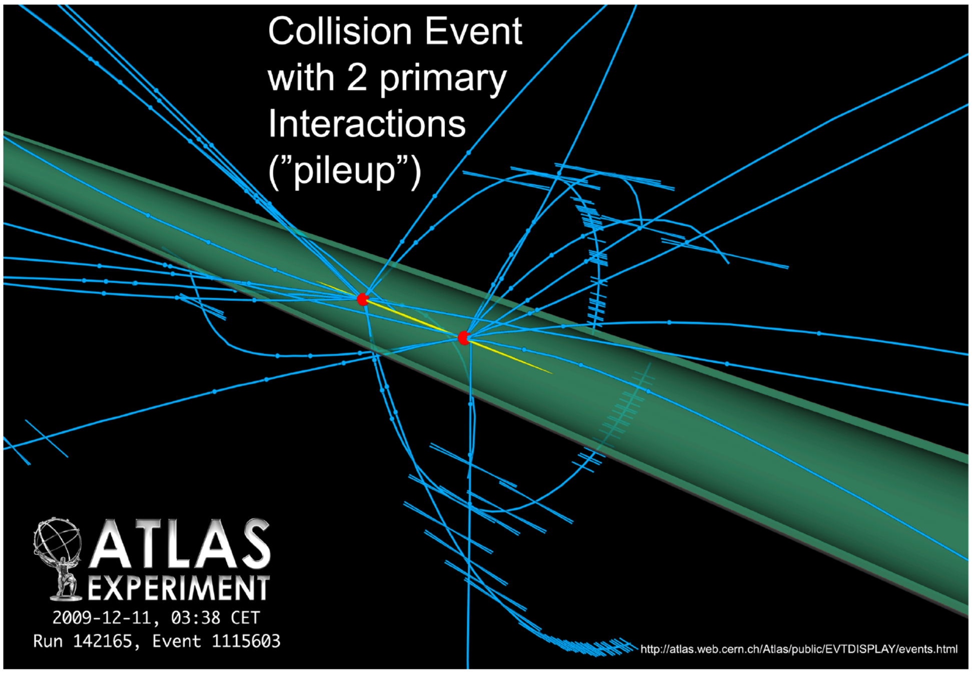



Figure 6.3: ATLANTIS event display of pileup event.

The vertex reconstruction algorithm used for this analysis can reconstruct more than one vertex per event, therefore these additional vertices are used to estimate the magnitude of this effect. We aim to reject events containing pileup. While the statistical gain in keeping the pileup is small, the pileup events could introduce bias, particularly in the tails of the measured distributions in particular in the dependence on the number of charged particles. We therefore decided to remove as many such events as possible and to estimate any potential residual effects in systematic error.

In order to illustrate the effect of these multiple interactions, Figures 6.4 shows the two dimensional distribution of the number of tracks at the vertex for the first v*s*. the second vertex for the events that fail our pileup removal cut. This illustrates that we do indeed see pileup events in our detector for 7 TeV run.

Most events have few tracks associated to the second vertex. In addition, similar vertices are observed in simulation, which contains no pileup. The removal of such fake pileup events would lead to a significant bias in our data sample.

In this analysis events that have a second vertex with more than three tracks are rejected. If the second vertex has three or fewer tracks, all tracks from the event that pass our selection are kept.

The data are not corrected for the removed events. In total about 500 events are removed by the pileup cut, which corresponds to 0.1% of our data sample, consistent with

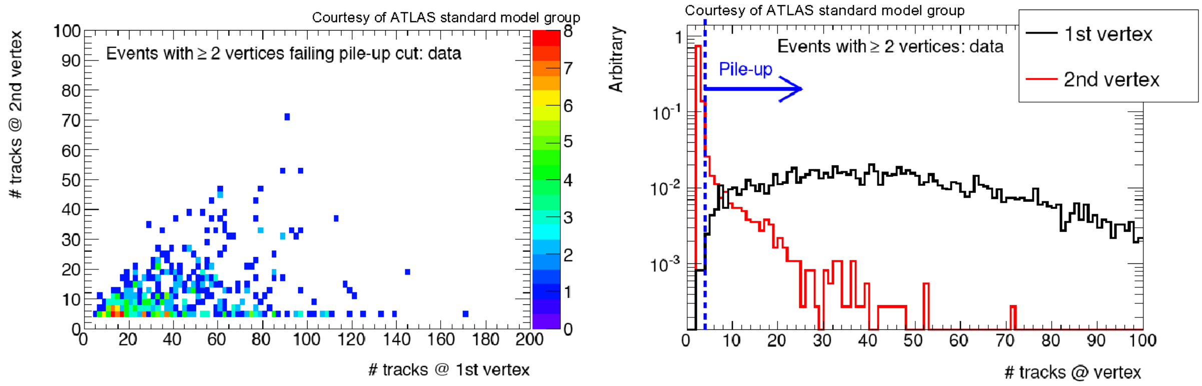



Figure 6.4: (left) The correlation between the number of tracks in the first and the second vertices. (right) The number of tracks at the first (black curve) and second (red) vertex. The pileup removal cut (above three tracks in the second vertex) is indicated by the blue dashed vertical line.

expectations.

Due to the nature of the vertex reconstruction algorithm, one expects to find, in 1% of events, additional vertices that either originate from the reconstruction of the decay vertex of secondary particles or are simply fake vertices. The left Figure 6.5 shows the number of events per LB after various stages including the pileup veto of our selection process at 7 TeV.

The number of events per LB when we require the vertex cut (dark green) plus the pileup (green) veto removing the second vertex decreases by about 1% as expected. It is shown in the right plot of Figure 6.5.

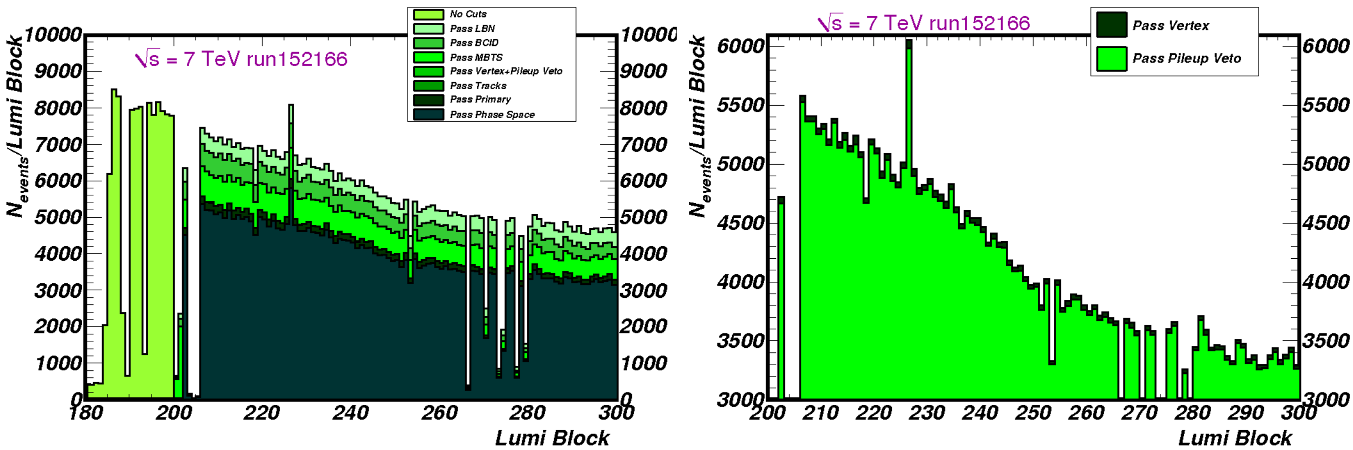



Figure 6.5: (Left) the number of events per LB after various stages including the pileup veto in our selection process at 7 TeV. (Right) the number of events per LB when we require the vertex cut (dark green) plus the pileup veto removing the second vertex (green).

The following different sources of systematic uncertainty are considered (further details in [27]).

- Events that are not pileup but are removed by the cut. This number is estimated to be $\sim 0.03\%$. The probability of having a fake split vertex depends on the number of tracks. It is below 0.1% for events with fewer than 30 tracks and increases to at most 1% for the highest bin.

- Events that are pileup but not removed by the cut. This number is estimated to be $\sim 0.01\%$.

- Events that are pileup but are reconstructed as a single vertex. This is estimated to be $\sim 0.01\%$.

All but the probability of a fake split vertex are derived from data and rely on generating pseudo experiments based on observed distributions. The probability of a fake split vertex is estimated by MC using a sample which does not contain pileup. The effect on the tails of the $n_{Sel}$ distribution were investigated. The fractional change in the number of events in each bin in $n_{Sel}$ is found to be less than 1% for all bins except for the last one, which is estimated to be $\sim 6\%$ with large uncertainties due to the lack of statistics in that bin. The effect on those tails due to the other two sources of systematic is smaller.

In conclusion, the pileup studies all show maximum systematic uncertainties well within the statistical uncertainty and the systematic uncertainty due to the track reconstruction efficiency, that will be explained in the next Section.

## 6.3 Track reconstruction efficiency

The quantity of interest for physics measurements is the efficiency to reconstruct a charged particle produced in the primary interaction. Inefficiencies arise from failures of the pattern recognition, failures of found tracks to pass the analysis selection cuts and cases where the particle undergoes an interaction in the detector with a significant change in momentum such that it becomes untrackable.

In this Section we measure the product of this efficiency directly in simulation by employing a cone method for matching the truth primary particles to the reconstructed tracks [18]. The systematic uncertainties associated to the matching are also discussed in Section 9.3.

### 6.3.1 Primary particle efficiency

The track reconstruction efficiency in each bin of the $p_T$-$\eta$ acceptance is determined from MC, justified by excellent agreement with data. Extensive comparisons between data and simulation are performed to establish that the simulation of the silicon detectors describes the data to a high level of accuracy. The observed discrepancy between data and simulation of the number of TRT hits on track has a negligible impact on track reconstruction efficiency and resolution.

As defined in Section 5.1.6, a secondary particle is defined as a particle which originates from hadronic interactions in the material of the detector, or from the decay chain of particles with lifetime greater than $3 \times 10^{-11}$ s. A particle is defined as primary otherwise.

From the practical point of view, particles created by Geant4 are considered as secondaries. These particles are recognised in the MC truth record as particles generated with a zero barcode or a barcode greater that 200.000.

The track reconstruction efficiency is determined, after full event selection, by dividing the distributions of all primary particles matched to a reconstructed track passing the selection cuts, by the corresponding distribution for all generated primary particles in the phase space:

$$\epsilon_{\mathrm{bin}}(p_{\mathrm{T}}, \eta) \; = \; \frac{n_{\mathrm{rec}}^{\mathrm{matched}}(p_{\mathrm{T}}, \eta)}{n_{\mathrm{gen}}(p_{\mathrm{T}}, \eta)}$$

To determine if a generated particle is reconstructed, a matching criteria needs to be defined to relate reconstructed tracks back to primaries. Few methods are studied in detail in [18] and [27], in this analysis we use the cone based matching. One other possible matching criteria, called hit-based matching, is based on comparing the number of hits on the track in each sub-detector to the number of hits produced by the true particle. The set of tracks matched using the hit-based matching, indicating good agreement between the two matching methods [18].

The cone based matching uses the difference in angles between the reconstructed track and the true primary to make the match. This is defined by:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{6.3}$$

where $\Delta\phi$ ($\Delta\eta$) are the differences between the truth track and the reconstructed track in $\phi$ ($\eta$). A match is considered successful if a reconstructed track is found with $\Delta R$ below a defined cut. If more than one reconstructed track is found in the cone around the truth particle, only the track with the smallest $\Delta R$ are matched. The parameters of the true primary are estimated at the perigee, while the parameters of the reconstructed track are estimated at the primary vertex. This introduces a small bias, but has no measurable effect on the primary efficiency.

If $\Delta R$ matching is used, the decay product of a true primary is matched provided that the direction of the reconstructed track is similar to that of the true primary. A cut of $\Delta R_{min} < 0.05$ is applied on the value of the minimal $\Delta R$. A track is considered as a fake track if it fails this matching procedure.

The resulting reconstruction efficiency as a function of $p_T$ integrated over $\eta$ is shown in Figure 6.6 for 900 GeV data sample (left) and 7 TeV data sample (right).

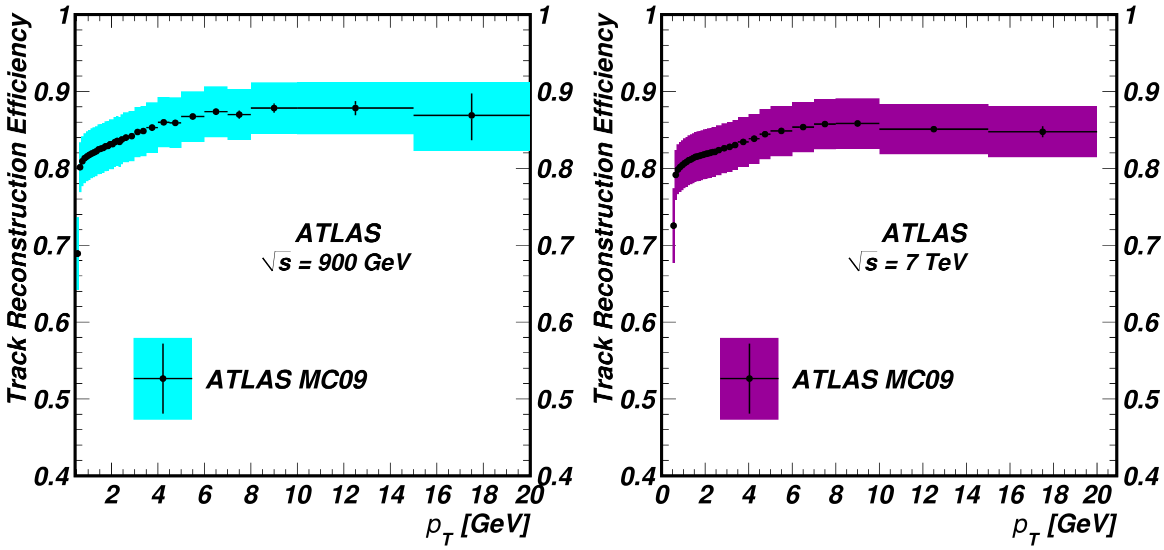



Figure 6.6: Track reconstruction efficiency as a function of $p_T$ for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The vertical bars represent the statistical uncertainty, while the shaded areas represent the statistical and systematic uncertainties added in quadrature.

The drop to about 70% for $p_T < 600\ MeV$ is an artefact of the $p_T$ cut at the pattern-recognition level. The consequence is that some particles which are simulated with $p_T > 500\ MeV$ are reconstructed with momenta below the selection requirement. This effect reduces the number of selected tracks. The shape of the threshold is studied in data and simulation and a systematic uncertainty of 5% is assigned to the first $p_T$ bin [18].

Relative to the published results we improve a little bit the track reconstruction settings. The only improvement affecting the track reconstruction efficiency is the lowering of the threshold of the track reconstruction algorithm used for 7 TeV analysis from 500 MeV to 100 MeV; this removes the turn-on effect due to the algorithm cut being at the same $p_T$ as the analysis cut and thus increases the track reconstruction efficiency in that $p_T$ region.

The resulting reconstruction efficiency as a function of $\eta$ is shown in Figure 6.7 for 900 GeV data sample (left) and 7 TeV data sample (right).

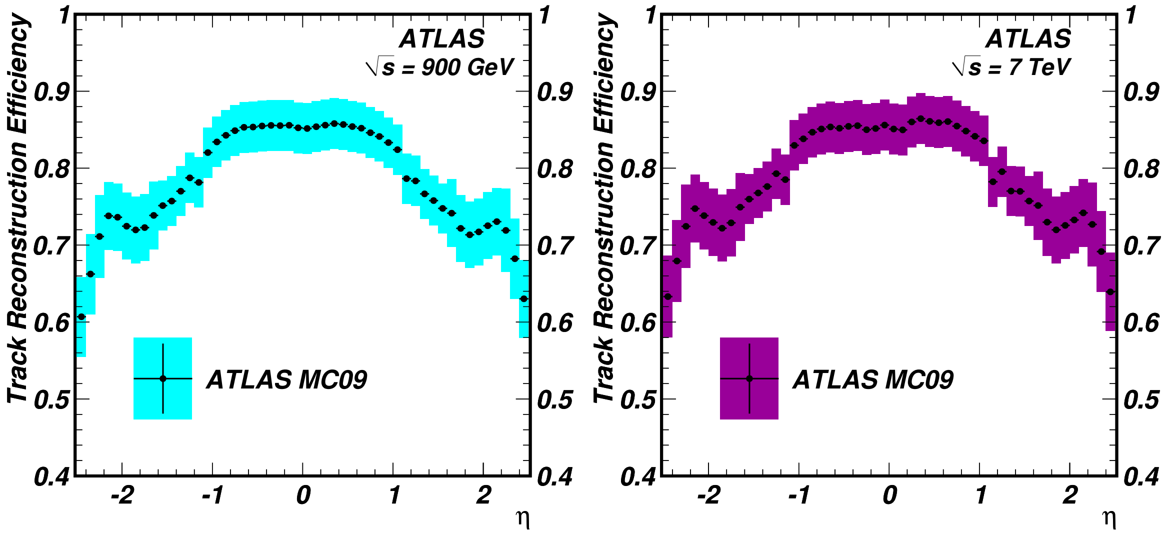



Figure 6.7: Track reconstruction efficiency as a function of $\eta$ for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The vertical bars represent the statistical uncertainty, while the shaded areas represent the statistical and systematic uncertainties added in quadrature.

The reduced track reconstruction efficiency in the region $|\eta| > 1$ is mainly due to the presence of more material in this region.

These inefficiencies include a 5% loss due to the track selection used in this analysis, approximately half of which is due to the silicon hit requirements and half to the impact parameter requirements.

The efficiency used in the corrections is parametrised in both $\eta$ and $p_T$.

### Tracking algorithm

Several tracking algorithms are available to reconstruct tracks in the ATLAS ID:

- Inside-out tracking: the default tracking algorithm. It reconstructs tracks seeded in the inner layers of the silicon detectors. Silicon tracks are reconstructed using at least seven silicon hits with a transverse momentum threshold of $p_T > 500$ MeV. Silicon track segments are then extrapolated to the TRT. A TRT extension is included in the final fit of the track if at least nine TRT hits are found and associated to the track.

- Back-tracking: this algorithm reconstructs tracks seeded in the TRT with a transverse momentum threshold of $p_T > 500$ MeV and then extrapolates them towards the silicon detectors. The back-tracking runs after the previous algorithm and uses only hits that were not yet used.

- Low momentum tracking: this is an inside-out tracking algorithm that runs with a lower transverse momentum threshold and a dedicated seeding algorithm to reconstruct tracks with $p_T > 150$ MeV. The minimal number of silicon hits requirement for the low momentum tracking is relaxed to require at least five hits. The low momentum tracking runs on the remaining hits after running the previous tracking algorithms.

- TRT standalone tracking: this algorithm reconstructs tracks in the TRT without attempting a silicon extension. It aims to reconstruct tracks that originate mainly from photon conversions and material interactions inside the TRT volume.

For this analysis, only tracks reconstructed with the standard inside-out tracking are used. At this early stage of the experiment, this algorithm is the most mature and the most studied. It will take more time to understand and validate the performance of the low momentum tracking.

## 6.4 Summary

Since the first pp collisions were observed in ATLAS on the 29th of November 2009, a number of studies of triggering, vertexing and tracking performance and efficiencies were performed. This chapter has focused on studies of triggering, vertexing and tracking efficiencies relevant to minimum bias physics measurements.

The trigger efficiency is measured in a data sample of events that contain at least one good offline track and are selected on the control sample trigger. The efficiency of the trigger is very high and is consistent with flat as a function of $p_T$ and $\eta$ of the tracks.

The vertex reconstruction efficiency is again determined from the data, by taking the ratio of triggered events with a reconstructed vertex to the total number of triggered events.

The track reconstruction efficiency in each bin of the phase space is determined from simulation.

Due to the improvements to the track and vertex reconstruction after the published analysis at 900 GeV [18], in this work we show the 900 GeV dataset reanalyzed using the new configuration only because the two results agree well, to within the assessed systematic uncertainties. Next step (Chapter 8) is to correct the selected sample of inelastic events for detector effects.

# Chapter 7

# Background contributions

This chapter describes the studies carried out to estimate the contamination from backgrounds to the event sample used for the minimum bias analysis on 900 GeV and 7 TeV datasets. There are two possible sources of background events that can contaminate the selected sample: cosmic rays and beam induced background.

## 7.1 Cosmic rays events

Cosmic rays events can contaminate the selected sample. A limit on the fraction of cosmic ray events recorded by the MBTS trigger during data taking is determined from:

- cosmic ray studies
- the maximum number of proton bunches
- the central trigger processor clock width of 25 ns.

  The contribution from this background source is found to be smaller than $10^{-6}$.

## 7.2 Beam induced background events

Beam induced background events can be produced by proton collisions with upstream collimators or with residual particles inside the beam pipe.

### 7.2.1 Beam background at 900 GeV

During the 900 GeV data taking, there were BCIDs where one beam is going past the centre of the ATLAS detector but there is no other beam to collide with. We called these BCIDs "unpaired bunches" or "single beam". The MBTS trigger is used to select beam induced background events from unpaired proton bunch crossings.

The MBTS detector consists of two wheels of 16 scintillators, these detectors are described in details in Section 2.2.1. One wheel is placed on each side of the interaction point, outside the endcap tracking devices. The MBTS are the detectors used mainly to trigger the signal events but we can use them also to get useful timing information. If the particles come from the centre of the detector, one would expect the timing to be roughly the same on each side. On the other hand, if the particles are coming from outside ATLAS and are traversing the detector, then from side to side one would expect the particles to first hit one side then the other, with a significant difference in timing. This fact can be exploited to select beam background events.

**MBTS time difference definition**

For each of the 32 counters, we measure the time of the reconstructed signal with respect to the LHC clock, i.e. the time when the bunch crosses the detector center. For collision events one would expect all counters to have roughly the same time, whereas for particles traversing the detector one would expect the time to be much earlier for counters on one side than on the other.

For each pair of scintillators, one for each side, the timing difference is defined as the difference in time for hits above an energy threshold of 200 MeV. It is defined as the difference between the A and the C side; thus a large positive time difference means that the particle reached the A side later than the C side, and vice-versa for a large negative time difference. We denote this quantity as $\Delta_{MBTS\ Time}$. For cases where no hits are above threshold the timing for that scintillator is not taken into account.

The computation of the MBTS timing difference can be expressed as

$$\Delta_{MBTS\ Time,i} = MBTS_i^{\mathrm{A\,side}} - MBTS_i^{\mathrm{C\,side}} \tag{7.1}$$

where $MBTS_i$ is the time for the counters above threshold on a given side.

By comparing the distributions of $\Delta_{MBTS\ Time}$ obtained from the collision events (paired bunches) with those of the single beam sample (unpaired bunches), one can estimate background events contaminating the sample used for the minimum bias analysis.

Figure 7.1 shows the MBTS timing difference for all events at 900 GeV before (left) and after (right) applying the offline cuts. The plot shows the result in logarithmic scale to show the details of the small tails. The events from collision BCIDs are shown in green and the single beam BCIDs are shown in blue. The peaks at $\pm 20$ $ns$ correspond to the time of flight for the distance between the MBTS position in A and C sides (6 m).

By applying the analysis selection criteria to these events, an upper limit of $10^{-4}$ is determined for the fraction of beam induced background events within the selected sample.

### 7.2.2 Beam background at 7 TeV

We study, as for the 900 GeV data, the effect of the beam backgrounds on 7 TeV data sample. The same method is used; based on the MBTS timing difference between the two sides of the detector, the results from single beam and collision events are compared.

In this case, differently from the 900 GeV analysis, the single beam data is collected using dedicated MBTS_1_UNPAIRED trigger. The unpaired proton bunch crossing are saved in dedicated stream.

In total, about 8.000 events are collected , of which only a single event passed the final event selection. Even making a conservative estimate of twice as much background in paired v$s$. unpaired bunches, the background is well below $10^{-3}$ and therefore negligible.

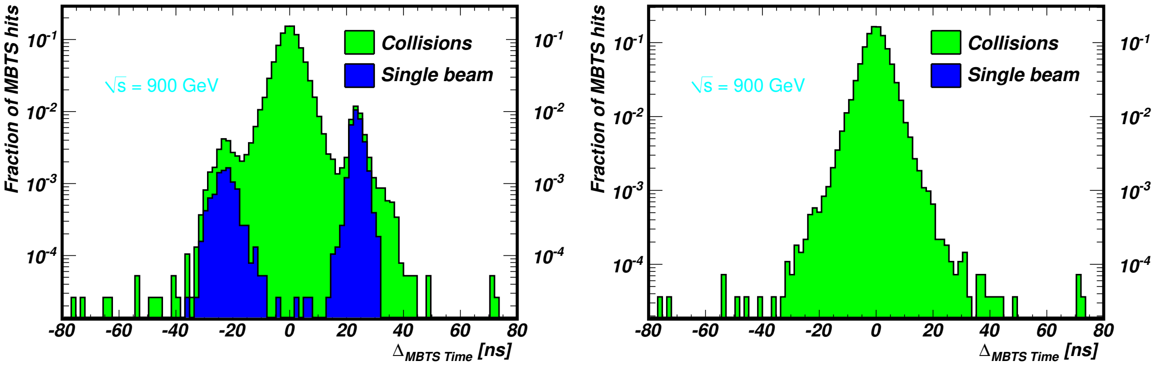



Figure 7.1: Timing difference for 900 GeV collisions between the A and C sides of the MBTS scintillators for collision events (green) and single beam events (blue) before applying the selection cuts (left) and after (right).

Figure 7.2 shows the MBTS time difference before and after offline cleaning cuts. Like for the 900 GeV data, the separation between the paired and unpaired can be clearly observed. The part of the distribution at $|\Delta_{MBTS\ Time}| > 60\ ns$ corresponds to a small fraction of the events where time reconstruction has failed. It has not impact on the estimate of the beam background rates.

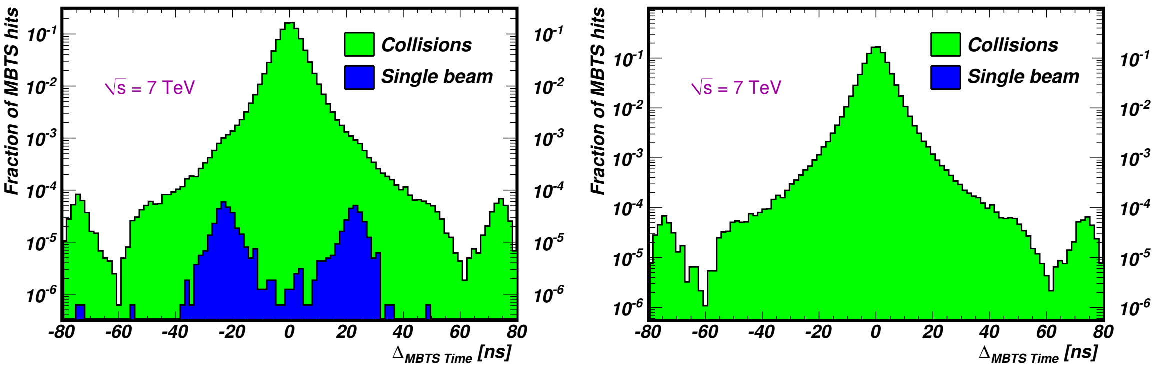



Figure 7.2: Timing difference for 7 TeV collisions between the A and C sides of the MBTS scintillators for collision events (green) and single beam events (blue) before applying the selection cuts (left) and after (right).

## 7.3 Potential background from fake tracks

To obtain the charged particle multiplicity, the charged track spectra are corrected for each step in event and track selection and reconstruction. The three major components are the trigger, the vertex and the track reconstruction efficiencies. In addition, the multiplicity must be corrected for the contribution from secondaries and fakes.

Defining a good track also raises the question of contamination through fake tracks, i.e. tracks that are formed from random hits that are not related to a true particle or to secondaries. MC is used to study this background. The techniques described in Section 6.3.1 are used to match reconstructed tracks to truth particles. This can then be used to measure fraction of fake tracks that are reconstructed. The fake rate is defined by the number of tracks that passed the track selection criteria and could not be matched to any true particle. For $\Delta R$ matching method with the cut at 0.05, the fake rate is less than $10^{-3}$.

## 7.4 Summary

This Section described the studies carried out to estimate the contamination from beam backgrounds to the event sample used for the minimum bias analysis on 900 GeV and 7 TeV data taken. During these data taking periods, we have some bunches circulating in the machine that did not collide in ATLAS. We thus use those events, known to be only background events to estimate the contamination in our signal. The difference in timing between the MBTS detectors on either side of the interaction point is used to investigate the shape and number of the background events in our sample. After applying the final event selection, one can estimate that the background contribution is less than $10^{-4}$ at 900 GeV and $10^{-3}$ at 7 TeV in terms of both events and tracks. This background can thus safely be neglected for the rest of the analysis. The fakes and the cosmic ray background are also found to be negligible for both analyses.

# Chapter 8

# Correction procedures

We measure primary tracks in minimum bias events produced in pp collisions at a centre-of-mass energy of 900 GeV and 7 TeV. Events have to pass a set of selection steps corresponding to the trigger, track reconstruction and selection, and primary vertex selection. Each of those selection steps causes inefficiencies that need to be corrected for in order to estimate the number of charged particles that are originally produced from the number of tracks measured in the selected events.

On top of these effects, particles from the primary collision can interact with the detector material to produce secondary charged tracks. What is seen in the detector is a set of tracks that contains the primary tracks, secondary tracks coming from a weakly decaying resonance ($K^0_S$, $\Lambda$, ...) or from material interactions, and fake tracks induced by the track reconstruction algorithm. These contaminations need to be subtracted from the final distributions.

We also correct for the fact that the detector and reconstruction algorithm have a certain resolution on the measured transverse momentum of the tracks, the detector $p_T$ is thus not identical to the true particle $p_T$.

In this Chapter we describe in detail how the above corrections are performed.

The aim is to obtain distributions that can be compared to theoretical models and other experiments without the need to simulate the experimental conditions of the ATLAS detector, as preferred by the theorists who could be using this data to tune MC models. These comparisons can only be performed in a common phase space. The phase space chosen for this analysis is at least one charged particle with a transverse momentum $p_T \geq 500$ MeV in a pseudorapidity range $\mid \eta \mid < 2.5$.

To avoid model dependencies of the result, no attempt is made to correct for events and tracks outside this phase space region.

## 8.1 Overview of corrections due to various inefficiencies

To obtain the four final distributions:

$$\frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ch}}}{\mathrm{d}\eta}, \quad \frac{1}{N_{\mathrm{ev}}} \cdot \frac{1}{2\pi p_{\mathrm{T}}} \cdot \frac{\mathrm{d}^2 N_{\mathrm{ch}}}{\mathrm{d}\eta \mathrm{d}p_{\mathrm{T}}}, \quad \frac{1}{N_{\mathrm{ev}}} \cdot \frac{\mathrm{d}N_{\mathrm{ev}}}{\mathrm{d}n_{\mathrm{ch}}} \quad \text{and} \quad \langle p_{\mathrm{T}} \rangle \text{ v}s.\ n_{\mathrm{ch}} \tag{8.1}$$

where $N_{\text{ev}}$ is the number of events with at least one charged particle inside the selected kinematic range, $N_{ch}$ is the total number of charged particles, $n_{ch}$ is the number of charged particles in an event and $\langle p_T \rangle$ is the mean $p_T$ for a given number of charged particles, at the particle level, one needs to apply two separate type of corrections for the track distributions. One type to correct the parameters of the measured tracks, $p_T$ and $\eta$, and the other to correct the event and track multiplicity, $N_{ev}$ and $n_{ch}$ respectively. While there is a large overlap in the procedure it is simpler to think of these corrections separately.

In both cases the effect of events lost due to the trigger and vertex requirements can be corrected using an event-by-event weight:

$$w_{\text{ev}}(n_{\text{Sel}}^{\text{BS}}) = \frac{1}{\epsilon_{\text{trig}}(n_{\text{Sel}}^{\text{BS}})} \cdot \frac{1}{\epsilon_{\text{vtx}}(n_{\text{Sel}}^{\text{BS}})}$$

where the efficiencies used in the correction terms are defined in Table 8.1.

| Correction | Notation |
|---|---|
| Trigger efficiency | $\epsilon_{trig}(n_{Sel}^{BS})$ |
| Vertexing efficiency | $\epsilon_{vtx}(n_{Sel}^{BS})$ |
| Track efficiency (binned) | $\epsilon_{bin}(p_T, \eta)$ |
| Event loss correction | $C_{tr}(\langle \epsilon_{bin}(p_T, \eta) \rangle, n_{Sel})$ |
| Fraction of secondary tracks | $f_{sec}(p_T)$ |
| Fraction of tracks outside kinematic range | $f_{okr}(p_T, \eta)$ |

Table 8.1: List of corrections applied to the raw data in order to get back to the particle level distributions.

The corrections to the $p_T$ and $\eta$ dependent distributions are obtained by applying weights $w_i(p_T, \eta, n_{Sel}^{\text{BS}})$ to every track $i$, where $w_i$ is given by

$$w_i(p_T, \eta, n_{Sel}^{\text{BS}}) = w_{\text{ev}}(n_{\text{Sel}}^{\text{BS}}) \cdot \frac{1}{\epsilon_{bin}(p_T, \eta)} \cdot (1 - f_{sec}(p_T)) \cdot (1 - f_{okr}(p_T, \eta)) \tag{8.2}$$

where the individual correction terms are again defined in Table 8.1. For the $dN_{ev}/dn_{ch}$ distribution the goal of the correction is to go from the measured number of tracks $n_{Sel}$ back to the number of particles $n_{ch}$. We can represent this transformation from tracks to charged primary particles as

$$n_{ch} = w_{\text{ev}}(n_{\text{Sel}}^{\text{BS}}) \cdot M_{ch,Sel} \cdot C_{tr}(\langle \epsilon_{bin}(p_T, \eta) \rangle, n_{Sel}) \cdot n_{Sel} \tag{8.3}$$

where $M_{ch,Sel}$ is the migration matrix to transform $n_{Sel}$ to $n_{ch}$, and $C_{tr}(\langle \epsilon_{bin}(p_T, \eta) \rangle, n_{Sel})$ is a term to correct for the event loss due to the track finding inefficiency and $\langle \epsilon_{bin}(p_T, \eta) \rangle$ is the mean value of the track efficiency, discussed below.

### 8.1.1 Inputs for the correction procedure

All correction terms, shown in Table 8.1, and their application are discussed in the following.

### The trigger efficiency $\epsilon_{trig}(n_{Sel}^{BS})$

The trigger efficiency is described in detail in Section 6.1. It is found to be flat with respect to the tracking variables but has a slight dependence on the number of tracks $n_{Sel}^{BS}$. It should be noted that this correction is derived before the vertex requirement and can thus not be parametrised with $n_{Sel}$, where a primary vertex is required. Therefore the correction is applied using weights parametrised to $n_{Sel}^{BS}$ instead of $n_{Sel}$.

### The primary vertexing efficiency $\epsilon_{vtx}(n_{Sel}^{BS})$

The primary vertexing efficiency is described in detail in Section 6.2. The vertex efficiency is defined in data as the fraction of events that pass the primary vertex selection after trigger. This definition assumes that the background is either negligible or behaves in an identical way to the signal. Beam backgrounds and fakes are found to be negligible.

The vertex efficiency strongly depends on the number of tracks in the event. It is found to be dependent on the $\eta$ of the track for events with $n_{Sel}^{BS} = 1$ track. This dependency is taken into account. The $\eta$ dependence for events with more tracks is found to be small and is thus neglected in the corrections procedure. On the other hand, the vertex requirement is found to not depend significantly on the track $p_T$.

The careful reader might have noticed that the trigger efficiency and the vertex reconstruction efficiency are measured in a slightly different variable space than the final event selection. This is motivated by the fact that one can not apply $d_0$ and $z_0$ cuts with respect to the primary vertex when there is no primary vertex.

Instead one can apply a cut to the $d_0$ with respect to the beam-spot. The way in which the efficiency is taken into account in the analysis is that for each event, one can compute both the $n_{Sel}$, used to fill the distributions, and the $n_{Sel}^{\mathrm{no}\,d_0\,z_0}$, used in the vertex efficiency correction.

The latter is always going to be equal or larger than the former definition of the number of tracks. As described above, the different corrections are applied on a track by track level (or event by event level) and each track is given a weight $w$ dependent on its properties. It is easy to show how the vertexing is applied by giving an example: an event $i$ has $n_{Sel} = n$ and $n_{Sel}^{\mathrm{no}\,d_0\,z_0} = m$. The vertex efficiency applied to all tracks in that event will be given by the efficiency for $m$ tracks, regardless of the value of $n$: $\epsilon_{vtx}(m, \eta)$. This takes into account the correlations between $n_{Sel}^{\mathrm{no}\,d_0\,z_0}$ and $n_{Sel}$ and is given directly by the data itself.

### Primary track reconstruction efficiency $\epsilon_{bin}(p_T, \eta)$

The primary track reconstruction efficiency, is calculated from MC in Section 6.3. It is determined, after full event selection, by dividing the distributions of all primary particles matched to a reconstructed track passing the selection cuts, by the corresponding distribution for all generated primary particles in the phase space:

$$\epsilon_{\mathrm{bin}}(p_{\mathrm{T}}, \eta) \;=\; \frac{n_{\mathrm{rec}}^{\mathrm{matched}}(p_{\mathrm{T}}, \eta)}{n_{\mathrm{gen}}(p_{\mathrm{T}}, \eta)}$$

This is done in events with a reconstructed primary vertex and at least one track passing the selection. The efficiency depends both on the $p_T$ and the $\eta$ of the tracks and thus a two dimensional parametrisation of the efficiency is required.

The correction procedures for trigger, vertex and tracking do not rely on the modelling of the underlying physics, meaning the corrections are either derived from data directly, or found to be independent of the underlying physics as in the case of the track efficiency.

The track efficiency is measured after our final event selection, which is not what one expect if one has already corrected for the trigger and vertex efficiencies separately, this is justified since it is found not to affect significantly the track efficiency [18].

**The migration matrix, $n_{Sel}$ to $n_{ch}$** For the distributions depending on $n_{ch}$ ($n_{ch}$ itself and $\langle p_T \rangle$ vs. $n_{ch}$) the track efficiency is applied in the form of a matrix $M_{ch,Sel}$ calculated from MC, and illustrated in figure 8.1. $M_{ch,Sel}$ determines for a given observed number of tracks, $n_{Sel}$, the probability that the event has $n_{ch}$ primary particles. The matrix is normalised such that the sum over all elements for a given $n_{Sel}$ is unity, ensuring the number of events remains constant during the variable transformation.

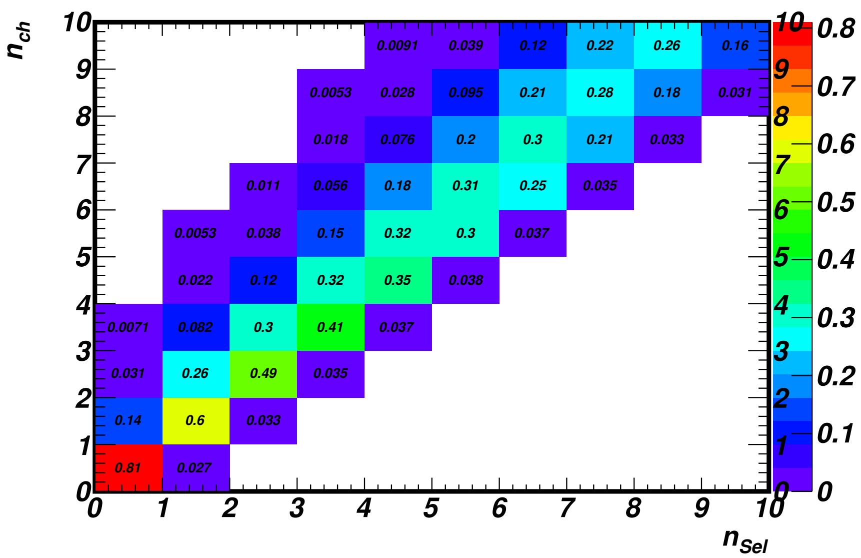



Figure 8.1: The plot shows the first $10 \times 10$ entries of the migration-matrix ($M_{ch,Sel}$), of lowest multiplicites at 900 GeV. The sum over all entries along the $n_{ch}$ axis for a given $n_{Sel}$ fixed to unity.

This matrix $M_{ch,Sel}$ depends on the MC model used. The difference on the integrated number of events when using a matrix derived from PHOJET and PYTHIA MC09 is of the order of 2% with a 5% difference in the first bin of $n_{Sel}$.

To reduce this MC dependence an iterative method is used for the matrix correction. In the first pass, one simply uses the raw MC to calculate the matrix. Then one compares the $n_{ch}$ distribution obtained in data after this correction to the one in the MC and re-weights the MC $n_{ch}$ distribution to that of the data. A new matrix is determined from this re-weighted MC and the procedure is repeated. After four iterations, there is little change in the final output.

### Event loss due to track inefficiency, $C_{tr}(\langle\epsilon_{bin}(p_T,\eta)\rangle, n_{Sel})$

The matrix $M_{ch,Sel}$ that is used to correct the number of tracks in the event back to the number of charged particles in the event can not take into account the effect of events that were simply not reconstructed.

The matrix method correction can not account for the loss of events due to track inefficiency because it only moves events from a particular value of $n_{Sel}$ to particular values of $n_{ch}$. If one considers the mean track efficiency to be $\langle\epsilon_{bin}(p_T,\eta)\rangle$, then $1-\langle\epsilon_{bin}(p_T,\eta)\rangle$ tracks are not reconstructed. Thus thinking in terms of the first bin in $n_{ch}$ where there is one primary charged particle: say one observes $n_{ch}$ events in that bin, we know there should have really been $n_{ch}/\langle\epsilon_{bin}(p_T,\eta)\rangle$ events.

To correct for this loss we apply the procedure described in [18].

The effect on the first bin (1 track) is 32%, 6% for 2 particles, 1% for 3 particles and negligible for higher charged particle multiplicities.

### Secondary tracks, $f_{sec}$

The fraction of secondary tracks $f_{sec}$ is estimated in Section 5.1.6. It is found that the rate of secondary tracks is described well in the MC. It is used as a $p_T$ dependent correction and is typically 2.20% at 900 GeV and 2.25% at 7 TeV.

### Migration of tracks from outside the kinematic range, $f_{okr}(p_T,\eta)$

Events that pass our offline event and track selection but that are outside the particle level phase space (outside the kinematic range) have to be removed. In other words we need to take in account the loosing events due the detector resolution at the boarder of the phase space.

Taking the $p_T$ as an example, we can have tracks that are reconstructed above 500 MeV but come from particles with a $p_T < 500$ MeV. A similar effect acts on the $\eta$ of the tracks. This will almost only affect the bins closest to the edge of our phase space cuts, i.e the first bin in $p_T$ and the edge bins in $\eta$. To remove such events we generate a 2D distribution $f_{okr}(p_T,\eta)$ of the affected event fraction from MC, and use it as a weight on the data [18]. The correction only affects the highest $\eta$ bins and behaves in a similar way to the lower $p_T$ region.

## 8.2 Effects due to the momentum scale and resolution

In this Section we focus on the transverse momentum scale and resolution for tracks in the low $p_T$ range, which are relevant for minimum bias measurements. To obtain corrected distributions of charged particles the scale and resolution of the $p_T$ have to be taken into account. The momentum resolution and momentum scale uncertainty is determined from simulation.

$K_S^0$ decays to two pions are used to quantify the $p_T$ resolution of low momentum tracks. The pion track $p_T$ is adjusted in scale and resolution in MC simulation in order to reproduce the measured $K_S^0$ mass and width in data. Dividing the dataset by the decay tracks pseudorapidity, allows to probe different detector regions. Furthermore, we separate the dataset depending on the track $p_T$. For low momentum particles, it is expected that the dominant contribution to the momentum resolution is due to multiple scattering, i.e. material effects [15], misalignment plays only a minor role due to the large curvature of the tracks. Hence, the determination of the momentum resolution in different $p_T$ regimes allows to test for and possibly disentangle detector misalignments and unaccounted material in the detector.

From the $p_T$ threshold shape at 500 MeV, the simulation is found to agree to 5%(±2%) in the resolution, which is further supported by the observed width of the $K_S^0$ mass distribution shown in the Figure 8.2.

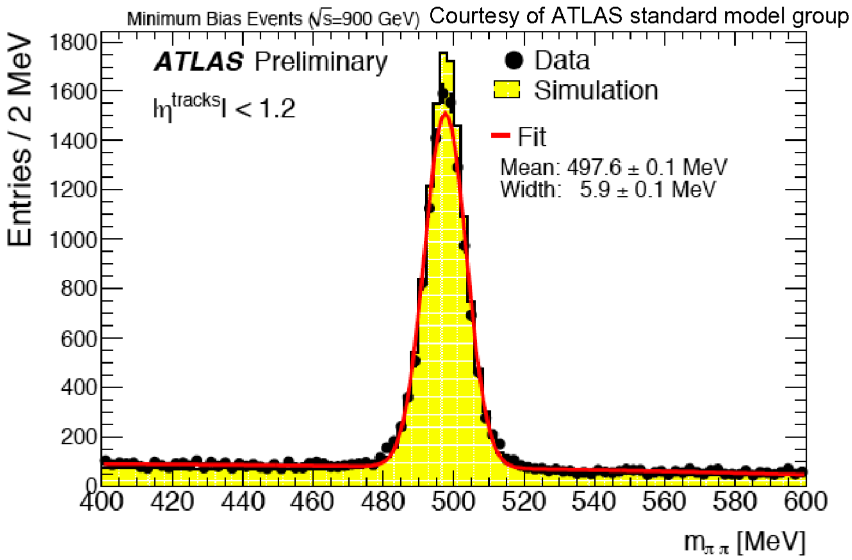



Figure 8.2: $K_S^0$ invariant mass distribution. In the plot one can see the barrel detector region with both tracks satisfy $|\eta| < 1.2$. The black circles are data, while the histograms show MC simulation (normalised to data). The red line is the line-shape function fitted to data.

Figure 8.2 shows the invariant mass distribution of two tracks from secondary vertex under pion assumption. The two tracks are required to have opposite charge. In addition, the distance of the reconstructed secondary vertex to the primary vertex has to be larger than 0.2 mm.

The scale is determined from the observed value of the $K_S^0$ mass which is found to be $497.6 \pm 0.1$ $MeV$ in excellent agreement to its true value of $497.614 \pm 0.024$ $MeV$. Therefore the biases coming from the momentum scale are negligible.

The determination of the transverse momentum resolution is described in [18]. In summary, from the decays of $K_S^0 \to \pi\pi$ one can estimate the $p_T$ resolution in data and compare it to the MC simulation. The data are found to agree well with the predictions. We then use the MC to parametrise the momentum resolution as a function of the $p_T$ and $\eta$ of the track.

### 8.2.1 Average transverse momentum distribution

The method used for the $\langle p_T \rangle$ vs $n_{ch}$ distribution is inspired by the method used by CDF [6]. The correction is done in two steps.

First we correct from $\langle p_T^{reco} \rangle$ to $\langle p_T \rangle$ for a given $n_{Sel}$. Then one corrects the number of tracks back to particles.

$M_{ch,Sel}$ is used to correct from $n_{Sel}$ to $n_{ch}$, taking $\langle p_T \rangle \cdot n_{ch}$ as a weight for each event, and normalising the resulting distribution to the sum of $n_{ch}$. The value of $\langle p_T \rangle$ for bin $j$ of $n_{ch}$ will thus be given by

$$\langle p_T \rangle_i = \frac{\sum_j M_i^j \langle p_T \rangle_j n_{Sel,j}}{\sum_j M_i^j n_{Sel,j}} \tag{8.4}$$

This corresponds to the generalised weighted mean of $\langle p_T \rangle$. The first stage of correcting the mean transverse momentum moves the points down respect to the observed distribution ($\langle p_T^{reco} \rangle$ vs $n_{Sel}$). This happens because the track efficiency is lower for low $p_T$ tracks; one is thus more likely to lose low $p_T$ tracks than higher $p_T$ ones during reconstruction, which artificially increases the observed mean value. At this point $\langle p_T \rangle$ vs $n_{Sel}$ is shown. The second stage of the correction moves the points mostly horizontally when transforming from $n_{Sel}$ to $n_{ch}$ using the matrix, also due to the track efficiency.

Because the corrections are rather small, no big uncertainty from the used MC model is expected.

The correction for the $\langle p_T \rangle$ is studied using the different PYTHIA tunes and PHOJET. The corrections applied to the average transverse momentum $\langle p_T \rangle$ are very small. This leads to small errors on the $\langle p_T \rangle$ distribution of the order of 1% for small $n_{ch}$. For high $n_{ch}$ the error is dominated by the track efficiency correction and low statistics in the matrix correction.

## 8.3 Summary

This Section described the correction procedure used to correct the data to obtain the distributions that can be compared to theoretical models.

To obtain the distribution of the number of charged particles $n_{ch}$ from the distribution of the number of measured tracks, first the trigger and vertex corrections are applied event-wise. Then the migration matrix $M_{ch,Sel}$ is applied. Finally the track loss correction factor $C_{Sel}$ is used.

# Chapter 9

# Systematic uncertainties

The total uncertainty on the final distributions is calculated as the sum in quadrature of the statistical uncertainty of the data and of the systematic uncertainties on the corrections applied to the raw distributions. Different sources of systematic uncertainties affecting one or more of the correction factors have been estimated.

In this section we explain the method used to estimate the effect of the systematic uncertainty on the final corrected distributions. Table 9.2 summarizes the experimental uncertainties from various sources considered in this analysis. For a thorough description of these uncertainties see the relevant references [18] and [27].

The effect of each individual source of uncertainty is propagated to the final result by shifting the corresponding quantity by one standard deviation and then repeating all the correction chain. The corrected distribution is then measured. The bin by bin difference with respect to the nominal value is considered as the systematic error. This process is repeated for each source of the systematic.

All sources of systematic uncertainties are assumed to be uncorrelated. The total systematic uncertainty is then calculated as the quadratic sum over all sources of uncertainty.

The systematic uncertainties from the analysis are propagated into the overall uncertainty on the corrected number of events ($N_{ev}$) and the charged particle multiplicity at $\eta = 0$.

The most relevant sources of systematic uncertainties are discussed in the following.

## 9.1 Uncertainties due to the trigger efficiency

The trigger efficiency dependence on the $p_T$ and $\eta$ distributions of the reconstructed tracks is found to be flat within the statistical uncertainties of the data recorded with the control trigger. The statistical uncertainty on this statement is conservatively taken as a systematic uncertainty of 0.1% on the overall trigger efficiency.

Since there is no vertex requirement in the data sample used to measure the trigger efficiency, it is not possible to make the same impact parameter cuts $(d_0, z_0)$ that are made on the final selected tracks. This means that the $n_{Sel}$ distribution is not exactly the same as that used in the analysis. In order to investigate how this may bias the measured efficiency, the selection based on impact parameters is varied. The efficiency of MBTS_1 has been analyzed for the following cases:

- no cut applied on $d_0^{PV}$ or $z_0^{PV}$
- cuts applied on $d_0^{PV}$ and $z_0^{PV}$ in events with an identified reconstructed vertex and no cut applied if no vertex is found
- cuts applied on $d_0^{PV}$ and $z_0^{PV}$ in events with a vertex, and on $d_0^{BS}$ in events without a vertex.

The latter is the default used to measure the efficiency. There is a small change in the efficiency in the low multiplicity bins. The difference is taken as a systematic uncertainty of 0.1%.

The correlation of the MBTS trigger with the control trigger is studied using the simulation[1)]. The resulting systematic uncertainty is found to affect only the case $n_{Sel}^{BS} = 1$ and amounts to 0.2%.

## 9.2 Vertex reconstruction efficiency

The vertex reconstruction efficiency and the corresponding uncertaintiy is used to correct the number of events measured in the data [18] [27]. Figure 6.2 shows that inefficiencies are observed up to the 4th bin and that the vertex reconstruction is fully efficient for higher number of selected tracks.

For the calculation of the systematic uncertainties on the measured efficiencies on the first bins of the distribution, studies of the following sources of the systematics are performed:

- Preselection of tracks for the vertex reconstruction and cuts on the quality of the reconstructed vertices. An estimate is obtained by studying the variation of the reconstruction efficiency between the data runs considered for the analysis.
- Contamination by beam background. An estimate is obtained by studying the events where only one bunch is crossing the ATLAS detector.
- Contamination by fake vertices. An estimate is obtained by studying the purity of the reconstructed primary vertices on the MC samples. After careful consideration it is decided that the vertices considered as fakes do not contribute to the systematic uncertainty. If the vertex is considered to be fake, i.e. contains majority of tracks coming from the secondary interactions and reconstruction errors, it still contributes to the final efficiency of the vertex reconstruction. There is thus no reason to assume a systematic contribution to the efficiency from these events.

In addition, a separate study is performed to estimate the dependencies of the vertex reconstruction efficiency in the first bins of the distribution on the parameters of the tracks

[1)]For 900 GeV data sample the default simulation of the MBTS trigger is not correct in the MC samples. The calibration constants used in the detector simulation to convert the energy deposited in the scintillator were wrong. For this reason an ad hoc simulation of the trigger is used, using the readout energy from the MBTS counters. The simulation overestimates the amount of energy deposited in a MBTS counter due to an incident particle.

selected for the analysis. No other sources of the systematic uncertainties are considered to be important for the present estimation of the efficiency.

The main and the only non negligible source of the systematic uncertainty is considered to be the contamination by beam background. All the other sources are found to have negligible contributions. The contribution of beam related backgrounds to the sample selected without a vertex requirement is estimated by using non colliding bunches. It is found to be 0.3% for $n_{Sel}^{BS} = 1$ and smaller than 0.1% for higher multiplicities, and it is assigned as a systematic uncertainty.

## 9.3 Track reconstruction efficiency

As the absolute tracking efficiency determined using the simulation, this section discusses the level to which the simulation describes the data.

The MC samples used consist of non-diffractive pp collisions produced using Pythia 6.42. Because the tracking efficiency is parametrized in $\eta$ and $p_T$, it does not depend strongly on the input sample, therefore only the non-diffractive sample is used. Two additional categories of simulation samples are used to quantify systematic uncertainties for the imperfect alignment and dead material of the detector.

### Truth primary definition

Two different techniques to associate generated particles to reconstructed tracks are studied: a cone matching algorithm and an evaluation of the fraction of simulated hits associated to a reconstructed track. The two non inclusive matching methods [18] obtain very similar values for the efficiency. The average difference between these two methods (of 0.4%) is assigned as a systematic uncertainty due to the matching method.

### Track selection

The selection efficiency for each cut is calculated as the ratio of the number of selected tracks after all cuts to the number of selected tracks after all cuts but the cut under study. This is defined in the following formula:

$$\epsilon_{cut} = \frac{n_{\mathrm{Sel}}(\mathrm{N\ cuts})}{n_{Sel}(\mathrm{N} - 1\ \mathrm{cuts})} \tag{9.1}$$

This technique has the advantage of not relying on any truth MC information, which means that the selection efficiency can be studied in both data and simulation. By comparing the results between data and simulation, the systematic uncertainties on efficiency of each cut are determined. For each cut, the efficiency is studied as a function of $p_T$, $\eta$ and $\phi$.

At 900 GeV the largest deviations between data and MC are observed by varying the $z_0 \sin\theta$ selection requirement, and by varying the constraint on the number of SCT hits. The simulation describes the selection efficiency observed in data very well in the central barrel where the systematic uncertainty is 0.4%. However, discrepancies are observed near $|\eta|$ of 2.5 where the systematic uncertainty reaches 2.2%

At 7 TeV the agreement between data and simulation is also excellent to 0.05% level, for the one pixel hit requirement. A 1% low efficiency for the impact parameter cuts is observed in data compared to simulation.

We assign 1% uncertainty due to understanding of the track selection efficiency in 7 TeV analysis, while uncertainty in 900 GeV analysis has small $\eta$ dependence as can be seen in the summary Table 9.1.

**Material**

Uncertainty on the material description of the detector impacts the tracking performance in two ways. First, the track fitting algorithm uses the expected material to estimate energy loss along the particle trajectory. Therefore uncertainty in the material degrades the resolution of the track parameters determined with respect to the global perigee. Second, the main source of tracking inefficiency is due to nuclear interactions of the particle with material within the ID. Uncertainty on the material results in an uncertainty on the rate of nuclear interactions. Furthermore the accuracy of the simulation to describe nuclear interactions depends on the ability of GEANT4 to model nuclear interactions.

There were not enough data to use (at the time of the analysis) photon conversions to constrain the material on the ID. Therefore a number of techniques have been explored to quantify the uncertainty on the tracking efficiency due to uncertainties on the material description.

Simulation samples with the nominal material increased by approximately 10% and 20% respectively are used to study the impact of additional material on the tracking efficiency. The fractional increase with respect to the total material is obtained by scaling the material of the support structure of the detector only.

The impact of the extra material in the tracking detectors is studied using:

- the tails of the impact parameter distributions ($d_0$ and $z_0$),
- the length of the tracks,
- the change in the reconstructed $K_S^0$ mass as a function of the decay radius,
- the direction and the momentum of the $K_S^0$.

The MC with nominal material is found to describe the data best. The data are found to be consistent to expectations within 10% except for perhaps in the very high $\eta$ regions, whereas the 20% increase is excluded in all cases. Integrated over the full acceptance, the total tracking efficiency decreases by 2.3% for 10% additional material and by 4.6% for 20% additional material. The decrease in efficiency depends on $\eta$ due to more material in the forward region.

The results also indicate that the nuclear interactions are well described by GEANT4 model in this kinematic regime. As a conservative systematic we assign a global systematic uncertainty on the tracking efficiency of 3%, which corresponds to approximately 20% additional material in the barrel region and approximately 10% additional material in the forward region. We maintain the same systematic for the 7 TeV analysis.

### ID alignment

The degree to which the detector is aligned can have a significant impact on the tracking efficiency. To fully estimate the systematic uncertainty from residual misalignment a simulation sample with the detector elements located according to the positions determined from the current alignment would be required. Then a comparison could be made between the tracking efficiency using the nominal alignment and with the best estimate of the detector positions.

Special MC samples are used to quantify systematic uncertainty due to a imperfectly aligned detector geometry. These samples have the detector module positions smeared to an alignment precision expected after either 1 or 100 days of operation.

The systematic effects of misalignment are studied by smearing simulation samples by the expected residual misalignment and by comparing the performance of two alignment algorithms on tracks reconstructed from the data. The change in the number of tracks is well below 1%, except for $\eta$ larger than 2 and large residual misalignment, when it becomes a 2% effect.

In conclusion, we keep the systematic effects of misalignment at the level of 1% for both analysis at 900 GeV and 7 TeV.

### SCT extension

Same comparisons between data and simulation show discrepancies of few % in the high pseudorapidity regions. This feature can be observed e.g. in the number of SCT hits per track, but is most pronounced in the extension efficiency of pixel tracklets. Pixel tracklets are tracks reconstructed only using the pixel detector. These require more hits in the pixel detector than combined tracks because at least three pixel hits are needed to constrain the number of degrees of freedom. The extension efficiency into the SCT is the probability to have found a combined track provided that a pixel tracklet has been reconstructed. It is sensitive to sub-system misalignment and to the material that is placed between the pixel and SCT detectors. The systematic errors due to these sources can not be easily assessed with any other methods discussed so far. We compare the efficiency to attach an SCT extension to these pixel tracklets in data and simulation. While excellent agreement is achieved in the central region, differences are observed at high $\eta$.

Several investigations of this effect, including studies of potential misalignment, excess in detector material and pattern recognition artifacts are ongoing but have not yet produced conclusive evidence for the source of this discrepancy. For the 900 GeV analysis, a conservative approach has been used, and the following relative systematic uncertainties on the tracking efficiency have been assigned due to discrepancies between data and simulation in which occur in the highest three bins in $\eta$:

- 10% ($2.4 \leq |\eta| < 2.5$)
- 6% ($2.3 \leq |\eta| < 2.4$)
- 2% ($2.2 \leq |\eta| < 2.3$).

The situation improves for the 7 TeV analysis:

- 6% ($2.2 \leq |\eta| < 2.5$)
- 2% ($1.6 \leq |\eta| < 2.2$).

**Low $p_T$**

For the 900 GeV analysis the track selection cut at $p_T > 500$ MeV in the track selection is identical to the internal cut used in the several stages of the pattern recognition. This makes the tracking efficiency at the low-$p_T$ edge sensitive to resolution and bias of the different pattern recognition stages.

As discussed in Section 6.3.1 the inside-out tracking is a four-stage process: a seed finding algorithm uses the space points found in the silicon detectors, a track candidate creation from the silicon seeds, a subsequent track cleaning and scoring in the ambiguity solving and finally the probe for extension into the TRT detector. The latter three components apply a requirement on the minimum $p_T$, and are thus sensitive to the momentum resolution and any bias to the momentum scale.

The turn-on characteristics can be studied by constructing the ratio of selected tracks within a specific $p_T$ bin to the subsequent bin. Given an exponential distribution, this ratio should be constant as is observed for $p_T > 800$ MeV [18]. Any deviation from an exponential distribution translates into an offset from this constant value. The magnitude of a negative offset provides an indication of the loss in reconstruction efficiency in the bin. In simulation, the plateau value is only reached for $p_T > 750$ MeV, while in data this occurs for $p_T > 600$ MeV. This indicates that there is a difference in the reconstruction efficiency between data and simulation in this momentum range. A comparison between data and simulation in the range of $0.5 < p_T < 0.6$ GeV of this ratio found a discrepancy of 2%.

An estimate is instead modeled using a toy MC by taking twice the total observed bias on the track candidates in simulation and assuming a 25% discrepancy on the $p_T$ resolution. This is pessimistic because there are no indications that either of these effects have discrepancies at this scale. This yields a relative difference in the efficiency of 6% between $p_T$ of 500 and 600 MeV; but no change is observed in any other $p_T$ bin. A relative systematic error of 5% has been assigned to the first bin in transverse momentum for the 900 GeV analysis. The low-$p_T$ systematic at 7 TeV is reported in the dedicated paragraph.

**Particle composition**

The tracking efficiency may vary depending on the particle type. Therefore, the final tracking efficiency is sensitive to the particle composition (the relative fraction of pions, kaons, protons, muons and electrons in the minimum bias sample). The tracking efficiency assumes that the composition is well described by the non-diffractive minimum bias simulation sample. The composition of primary particles within the analysis phase space as modeled by PYTHIA is as follows

- pions: $\sim$ 77.4%
- kaons: $\sim$ 14.9%

| **Systematic Uncertainty** | **Systematic at** $\sqrt{s} = 0.9\ TeV$ | **Systematic at** $\sqrt{s} = 7\ TeV$ |
|---|---|---|
| Truth Primary Definition | $\pm 0.4\%$ | $\pm 0.4\%$ |
| Track Selection | $\pm\, \mathbf{0.4}\%$ ($\lvert\eta\rvert < 2.3$)<br>$\pm 2.2\%$ ($\lvert\eta\rvert > 2.3$) | $\pm\mathbf{1}\%$ |
| Material | $\pm\mathbf{3}\%$ | $\pm\mathbf{3}\%$ |
| Alignment | $\pm 1\%$ | $\pm 1\%$ |
| SCT Extension | $\pm 10\%$ ($2.4 < \lvert\eta\rvert < 2.5$)<br>$\pm 6\%$ ($2.3 < \lvert\eta\rvert < 2.4$)<br>$\pm 2\%$ ($2.2 < \lvert\eta\rvert < 2.3$) | $\pm 6\%$ ($2.2 < \lvert\eta\rvert < 2.5$)<br>$\pm 4\%$ ($1.6 < \lvert\eta\rvert < 2.2$) |
| Low $p_T$ | $\pm 5\%$ ($0.5 < p_T < 0.6$ GeV) | $\pm 1\%$ ($0.5 < p_T < 0.6$ GeV) |
| Particle Composition | $\pm 0.2$ | $\pm 0.2$ |
| Total | **3.1%** ($p_T > 0.6$ GeV; $\lvert\eta\rvert < 2.3$) | **3.8%** ($p_T > 0.6$ GeV; $\lvert\eta\rvert = 0$)<br>3.9% ($0.5 < p_T < 0.6$ GeV; $\lvert\eta\rvert = 0$)<br>7.1% ($p_T > 0.6$ GeV; $2.4 < \lvert\eta\rvert < 2.5$) |

Table 9.1: Summary of systematic uncertainties on the tracking efficiency in 900 GeV and 7 TeV analyses.

- protons: $\sim$ 7.6%

The contribution from muons and electrons is negligible.

To assess the systematic uncertainty associated to this assumption the relative fraction of each particle type is varied in the following way: pions, kaons and protons varied by $\pm 10\%$, electrons and muons varied by a factor of 3 and 1/3. The variation range is obtained by comparing the particle composition between PYTHIA and PHOJET samples. Each particle type is varied independently and the change to the $\eta$ and $p_T$ distributions are determined. The maximum observed differences are added in quadrature for the variations of the different particle types. The systematic uncertainties obtained are below 0.2% for $p_T > 500$ MeV in both analyses.

Table 9.1 summarises the various contributions to the systematic uncertainties on the tracking efficiency for both analyses at 900 GeV and 7 TeV. The final tracking efficiency distributions including the systematic uncertainties (v$s$. $\eta$ and $p_T$) are shown in Figures 6.6 and 6.7. The largest systematic uncertainty is due to the uncertainty in the ammount of material in the detector compared to the simulation. There is also a large systematic at $\lvert\eta\rvert > 2.3$ due to differences observed between data and simulation in these regions. The uncertainty due to possible differences in the $p_T$ turn on region ($p_T < 600$ MeV) leads to a significant additional systematic uncertainty for tracks in these $p_T$ region. The largest uncertainty is on the pattern recognition efficiency below 1 GeV. The systematic uncertainty on the number of disabled pixels is not present since the simulation sample with the disabled pixel modules is available in both analysis at 900 GeV and 7 TeV.

**Systematic uncertainties on the track reconstruction efficiency for the 7 TeV analysis**

The method used to extract the systematic uncertainties on the track reconstruction efficiency at 7 TeV has remained unchanged, with the exception that, because changes in the tracking algorithm, there is no longer a systematic uncertainty associated to the fact that the algorithm and analysis cuts are both at 500 MeV.

The most significant changes are the improvements to track reconstruction at low momentum, which decrease the systematic uncertainty on the tracking efficiency, the use of more sophisticated vertex reconstruction algorithms and the removal of pileup (see Section 6.3).

Then the systematic uncertainty in the first bin is not applicable any more to the 7 TeV analysis due to the improvements the track reconstruction algorithm. However, due to the cut on the $p_T$ of the primary particle, the efficiency in the first bin is sensitive to the $p_T$ resolution. Again a toy MC is used to determine that a conservative uncertainty of 10% on the $p_T$ resolution changed the efficiency by 1%. The change is insensitive to the $p_T$ spectrum used in the toy MC. This is taken as the systematic uncertainty on the track reconstruction efficiency in the first $p_T$ bin due to the momentum resolution.

The systematic uncertainties due to the application of the various track reconstruction cuts is obtained, as for the 900 GeV analysis, by removing each cut in turn and comparing the observed effect between data and MC. The difference between them is taken as the systematic associated to that cut. The total systematic associated to the track selection is taken as 1% within all the acceptance.

The remaining sources of systematic are not re-evaluated for this analysis. Table 9.1 summarizes again the various contributions to the track reconstruction uncertainty and other sources of systematics uncertainties for the 7 TeV analysis. The truth primary definition is the systematic associated to the criteria used to match a hadron level particle with a reconstructed track. The SCT extension efficiency is derived by comparing the fraction of pixel tracks that have an reconstructed extension in the SCT between data and MC [18].

In conclusion in both analysis, an overall relative systematic uncertainty of 3.0% to 4% is assigned to the track reconstruction efficiency for most of the kinematic range of this measurement, a larger value is assigned to the highest $\eta$ and to the lowest $p_T$ bins.

## 9.4 Momentum scale and resolution

To obtain corrected distributions of charged particles, the scale and resolution uncertainties in the reconstructed $p_T$ and $\eta$ of the selected tracks have to be taken into account. The ID momentum resolution is taken from MC as a function of $p_T$ and $\eta$. It is found to vary between 1.5% and 5% in the range relevant to this analysis.

For low $p_T$ tracks, below 2 GeV, the uncertainty is taken as the difference between the resolution function obtained using the nominal MC and the sample generated with an addition of 10% material. Above 2 GeV, the difference between the resolution obtained from nominal and day one alignment is taken as the uncertainty.

The uncertainty is estimated by comparing with MC samples with a uniform scaling of 10% additional material at low $p_T$ and with large misalignments at higher $p_T$. Studies of the width of the mass peak for reconstructed $K_S^0$ candidates in the data show that these assumptions are conservative. The reconstructed momentum scale is checked by comparing the measured value of the $K_S^0$ mass to the MC.

The mass of the $K_S^0$ is sensitive to the amount and distribution of material through electromagnetic dE/dx energy loss rather than through nuclear interactions. Therefore, it is sensitive to the number of radiation lengths rather than the number of interaction lengths. The conversion between the two depends on the $A$ and $Z$ of the material traversed.

Comparisons of the data to the MC samples indicates that the data falls between the nominal geometry and the +10% sample. This $K_S^0$ study is presented in more detail in [18]. In both analysis the systematic uncertainties from both the momentum resolution and scale are found to give a negligible contribution to the final results.

## 9.5 Correction procedure (different Monte Carlo tunes)

For the corrections derived from MC, the statistical uncertainty of the MC sample needs to be taken into account. This is derived simply as $\sqrt{N_{MC}}$ where $N_{MC}$ is the number of MC events (or tracks, where relevant) in a particular bin.

A source of uncertainty not described by the individual efficiency measurements is the uncertainty due to the matrix used to correct the $n_{Sel}$ to the $n_{ch}$ distribution already illustrated in Section 8.1.1. After using the iterative unfolding procedure, the uncertainty due to the specific MC model used to generate this matrix is greatly reduced. The residual systematic is taken as the difference in the unfolded distributions when using PHOJET vs the default, MC09 tune of PYTHIA. The uncertainty on this matrix, after the iterative procedure is applied, is taken to be 3% up to $n_{ch}$ of 24 and 10% above [18].

Another source of systematic uncertainty affecting this matrix is the value of the track efficiency. We randomly move events in the matrix according to the tracking efficiency shift taken from Table 9.1, repeating once for a downwards shift and once for an upwards shift. We then re-calculate the final distributions and the difference with respect to the nominal matrix is taken as the systematic. The shifted matrix is compared to the original one and the difference taken as a systematic, found to be of the order of 4% in the first bin.

## 9.6 Fraction of secondaries

The fraction of secondaries is determined as discussed in Chapter 7. The associated systematic uncertainty is estimated by varying the range of the impact parameter distribution that is used to normalise the MC, and by fitting separate distributions for weak decays and material interactions. The systematic uncertainty includes a small contribution due to the dependence of this correction. The total uncertainty is 0.1% in both analyses.

## 9.7 Systematic uncertainties on the number of events, $N_{ev}$ and on the charge particle density $(\frac{1}{N_{ev}})(\frac{dN_{ch}}{d\eta})$ at $\eta = 0$

The systematic uncertainty on the normalisation and on the number of charged particles are treated separately. In each of these two groups the systematic uncertainties are added in quadrature. These are then combined taking into account their anticorrelation and are propagated to the final distributions.

In Table 9.2 we present the summary of systematic uncertainties on the number of events, $N_{\mathrm{ev}}$, and on the charged particle density $(1/N_{\mathrm{ev}})\cdot(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)$ at $\eta = 0$.

| **Systematic uncertainty on the number of events, $N_{\mathrm{ev}}$** | | |
|---|---|---|
| | $\sqrt{s} = 0.9\ TeV$ | $\sqrt{s} = 7\ TeV$ |
| Trigger efficiency | 0.2% | 0.2% |
| Vertex reconstruction efficiency | $<$ 0.1% | $<$ 0.1% |
| Track reconstruction efficiency | 1.1% | 0.8% |
| Different MC tunes | 0.4% | 0.4% |
| Total uncertainty on $N_{\mathrm{ev}}$ | 1.2% | 0.9% |
| **Systematic uncertainty on $(1/N_{\mathrm{ev}})\cdot(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)$ at $\eta = 0$** | | |
| Track reconstruction efficiency | 3.1% | 3.8% |
| Trigger and vertex efficiency | $<$ 0.1% | $<$ 0.1% |
| Secondary fraction | 0.1% | 0.1% |
| Total uncertainty on $N_{\mathrm{ev}}$ | $-1.2$% | $-0.9$% |
| Total uncertainty on $(1/N_{\mathrm{ev}})\cdot(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)$ at $\eta = 0$ | 1.9% | 2.9% |

Table 9.2: Summary of systematic uncertainties on the number of events, $N_{\mathrm{ev}}$, and on the charged particle density $(1/N_{\mathrm{ev}})\cdot(\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta)$ at $\eta = 0$. The uncertainty on $N_{\mathrm{ev}}$ is essentially anti-correlated with $\mathrm{d}N_{\mathrm{ch}}/\mathrm{d}\eta$. All sources of uncertainty are assumed to be uncorrelated. The correlation between $N_{ev}$ and the numerator are fully taken into account in the error calculation but are shown here separated out for illustration purposes.

The only significant systematic error on $N_{ev}$ is the correction for events which are lost due to track inefficiency which is of the order of 1% percent in both analyses.

For the charged particle density the main contribution to the total error is the correction to the track efficiency due to material and the correction for loss of events due to tracking inefficiency. Both effects are fully correlated and less track efficiency leads to more events in the low multiplicity bins. Another big contribution to the total error is the model dependence of the migration matrix. This error is reduced by the iterative unfolding procedure.

The trigger efficiency ($\leq 0.1\%$), the vertex efficiency (0.4%) and the correction for secondaries (0.1%) are small sources of systematics.

Since the track efficiency is anti-correlated to the loss of events, resulting in a change of $N_{ev}$, applying the $1/N_{ev}$ normalisation leads to a reduction of the flat systematic error by 1.2% at 900 GeV and 0.9% at 7 TeV.

# Chapter 10

# Results

In the previous Chapters we have outlined the procedure used to correct the data to obtain the distributions that can be compared to theoretical models. The effect of systematic uncertainties on the final distributions have been quantified.

We present now the first measurements of charge particle multiplicity distributions at $\sqrt{s} = 900$ GeV and $\sqrt{s} = 7$ TeV with ATLAS detector. Data are collected in 2009 and 2010 respectively, using the minimum bias trigger scintillators. The charged particle multiplicity, its dependence on transverse momentum and pseudorapidity, and the relationship between mean transverse momentum and charged particle multiplicity are measured for events with at least one charged particle within the kinematic range $|\eta| < 2.5$ and $p_T > 500$ MeV.

The analysis at 7 TeV [27] repeats as closely as possible the strategy used for the 900 GeV analysis [18]. However, improvements in track reconstruction and vertexing have been used. Therefore the the 900 GeV dataset is reanalysed, to confirm that this had no impact on the published results[1)].

The measurements are compared to MC models of pp collisions and to results from other experiments at the same centre-of-mass energy.

## 10.1 Inclusive inelastic distributions with minimal model dependent corrections

The final plots with all corrections applied and systematic uncertainties calculated are shown in this Section. For all the final plots, the data points are shown in black with the error bars representing the statistical uncertainty and the shaded area showing the total uncertainty on each bin. The data are presented as inclusive inelastic distributions with minimal model dependent corrections to facilitate the comparison with different models. The values of the ratio histograms refer to the bin centroids.

The charged particle pseudorapidity density is shown in Figure 10.1 for $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$ data samples. It is measured to be approximately flat in the range $|\eta| < 1.5$, with an average value of $1.333 \pm 0.003$ (*stat.*) $\pm$ 0.027 (*syst.*) and 2.418 $\pm$

[1)] In this work we only report the 900 GeV data reanalysed. Further details can be found in [27].

0.004 $(stat.) \pm 0.076$ $(syst.)$ charged particles per event and unit of pseudorapidity in the range $|\eta| < 0.2$ at 900 GeV and 7 TeV respectively.

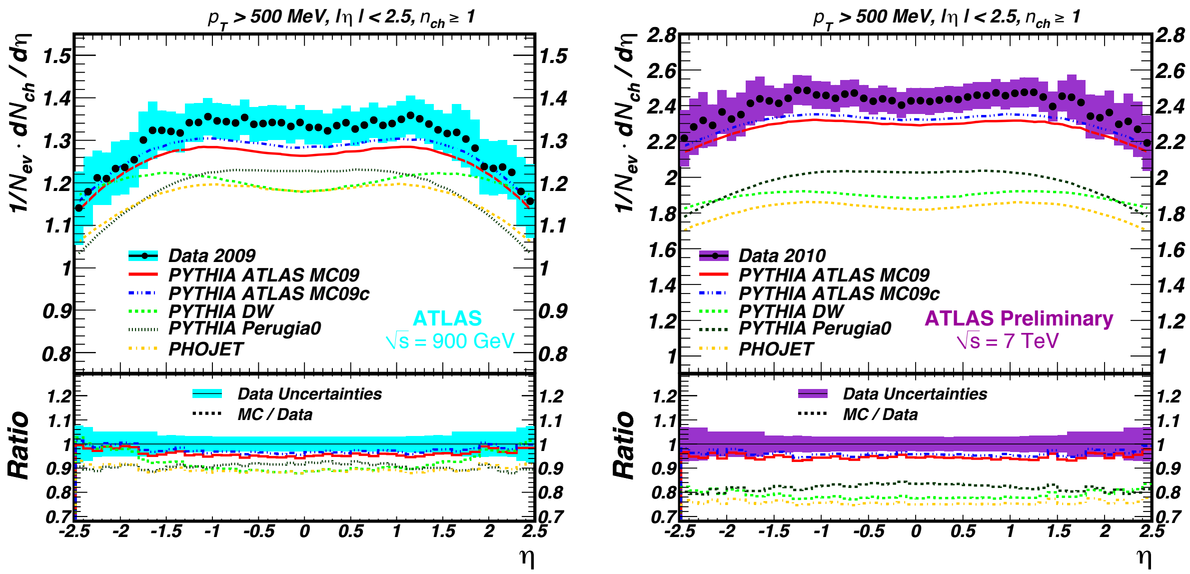



Figure 10.1: Charged particle multiplicity as a function of pseudorapidity ($\eta$) for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

The particle density is found to drop at higher values of $\eta$. All MC tunes discussed in this thesis are lower than the data by 5-15%, corresponding to approximately 1-4 standard deviations. The shapes of the models are approximately consistent with the data with the exception of PYTHIA DW tune. The distribution is best described by the ATLAS MC09c tune which has the same shape but is about 5% lower. None of the models are able to describe the data precisely over the whole $p_T$ spectrum.

The charged particle $p_T$ density is shown in Figures 10.2 and 10.3 in log-log and linear scales again for the 900 GeV data sample in the left plot and for the 7 TeV data sample in the right plot. The data distribution span over eight orders of magnitude. The spectrum is well described by ATLAS MC09 tune up to 2 GeV, which though predicts a significantly harder spectrum at higher $p_T$. It is interesting to note that all PYTHIA tunes based on the $p_T$ ordered shower (MC09, MC09c, Perugia0 tunes of PYTHIA) have a similar shape and a harder spectrum, while the tune based on the virtuality ordered shower (DW tune of PYTHIA and PHOJET) reproduces the full spectrum within 20%.

The multiplicity distribution as a function of $n_{ch}$ is shown in Figures 10.4 and 10.5 in linear-log and linear scales again for the 900 GeV data sample in the left plot and for the 7 TeV data sample in the right plot. The distribution is reasonably well described by the ATLAS MC09 PYTHIA tune for $n_{ch} > 5$ up to $n_{ch} = 50$. The excess of the model over data at lower $n_{ch}$ is highly influenced by the modelling of diffractive events in the MC

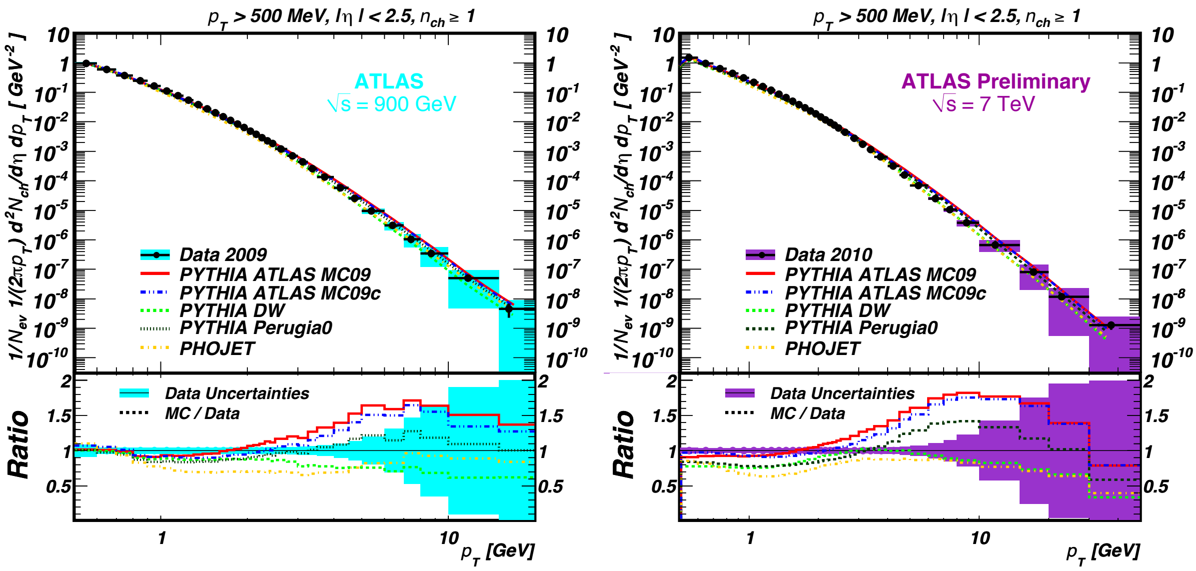



Figure 10.2: Charged particle multiplicity as a function of the transverse momentum ($p_T$) for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

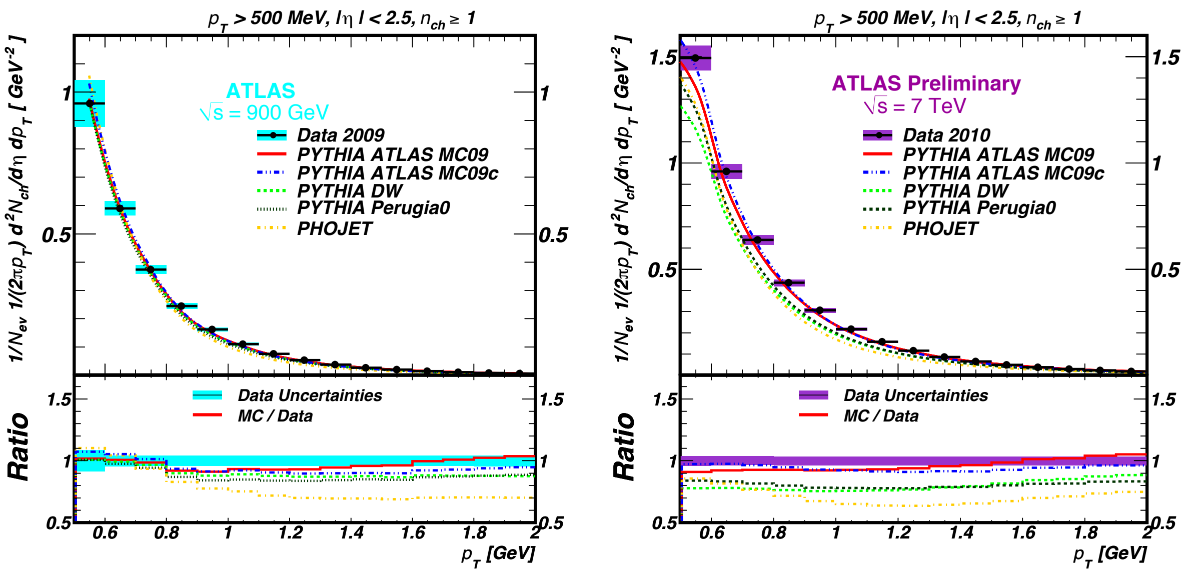



Figure 10.3: Charged particle multiplicity as a function of the transverse momentum ($p_T$) in linear scale for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

predictions, while the fraction of events with $n_{ch}$ above 10 is consistently lower than in the data. The net effect is that the integral number of charged particles predicted by the models are below that of the data (Figure 10.1).

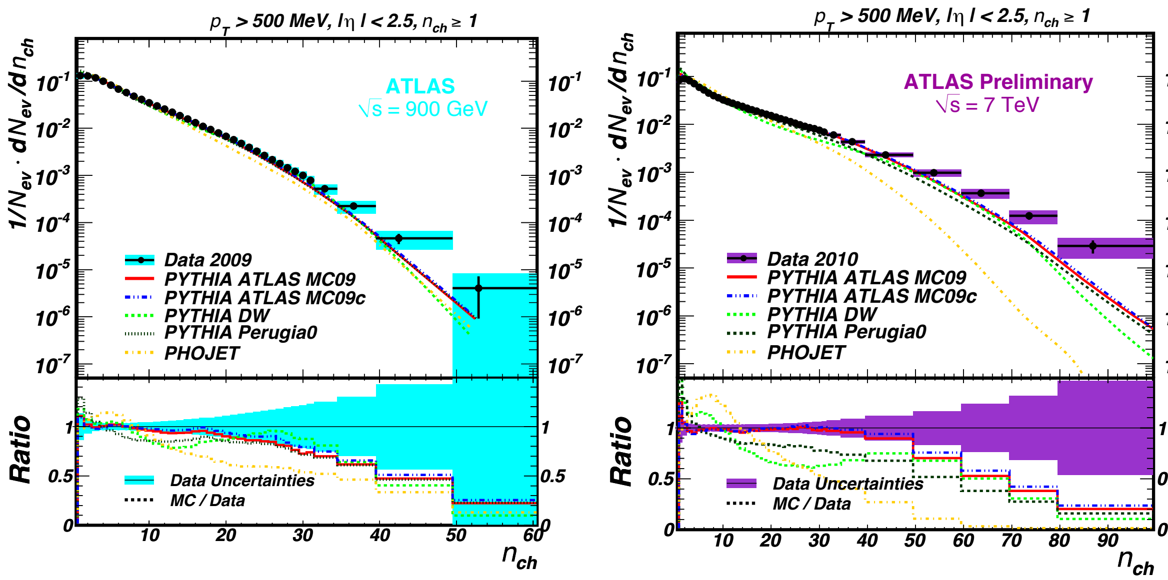



Figure 10.4: Charged particle multiplicity distribution for $\sqrt{s}$ = 900 $GeV$ in the left plot and $\sqrt{s}$ = 7 $TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

At 7 TeV the multiplicities of events with $n_{ch} > 50$ is higher in the data than in the PYTHIA tunes by about 70%. PHOJET fails to describe this distribution above $n_{ch} = 30$.

The average $p_T$ as a function of $n_{ch}$ is presented in Figure 10.6 for 900 GeV data sample on the left and for 7 TeV data sample on the right.

The average $p_T$ as a function of $n_{ch}$ increases with increasing $n_{ch}$ and a change of slope is observed around $n_{ch} \sim 10$. This behaviour was already observed by the CDF experiment in $p\bar{p}$ collisions at 1.96 TeV [6]. A recently published CDF measurement of the mean transverse momentum $\langle p_T \rangle$ as a function of the charged particle multiplicity $n_{ch}$ [6] shows that this distribution is not well described by the MC09 ATLAS tune. As shown in Figure 10.6, the model overshoots the data at high $n_{ch}$ while the Perugia0 tunes which used these data describes it very well, further details in [9], [10].

For 900 GeV data sample the Perugia0 parameterization, which is tuned using CDF minimum bias data at 1.96 TeV, describes the data well. The other models fail to describe the data below $n_{ch}$ 25, with the exception of the PYTHIA-MC09c tune.

For 7 TeV data sample this variable is relatively well described by the MC. The distribution is described by PHOJET within 3% up to $n_{ch}$ = 60 where the model shows a sudden rise which is not seen in the data. The PYTHIA tunes, Perugia0 and ATLAS MC09c describe it within 6% up to $n_{ch}$ =70.

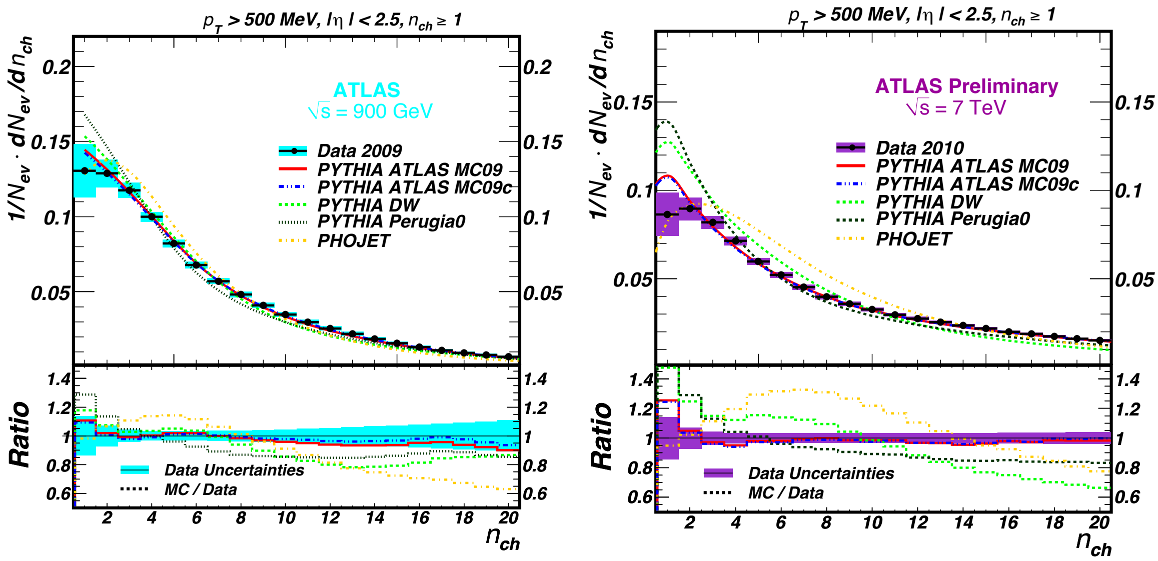



Figure 10.5: Charged particle multiplicity distribution in linear scale for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

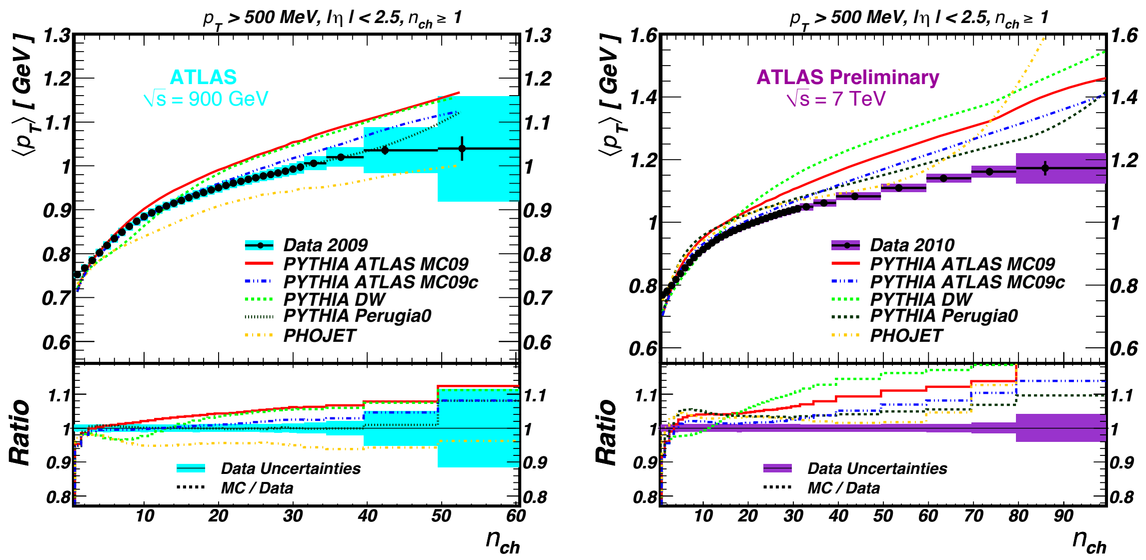



Figure 10.6: The average transverse momentum as a function of the number of charged particles in the event for $\sqrt{s} = 900\ GeV$ in the left plot and $\sqrt{s} = 7\ TeV$ in the right plot. The dots represent the data and the curves the predictions from different MC models. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

Figure 10.7 shows the comparison of the four observables comparing the published 900 GeV data to the 7 TeV data.

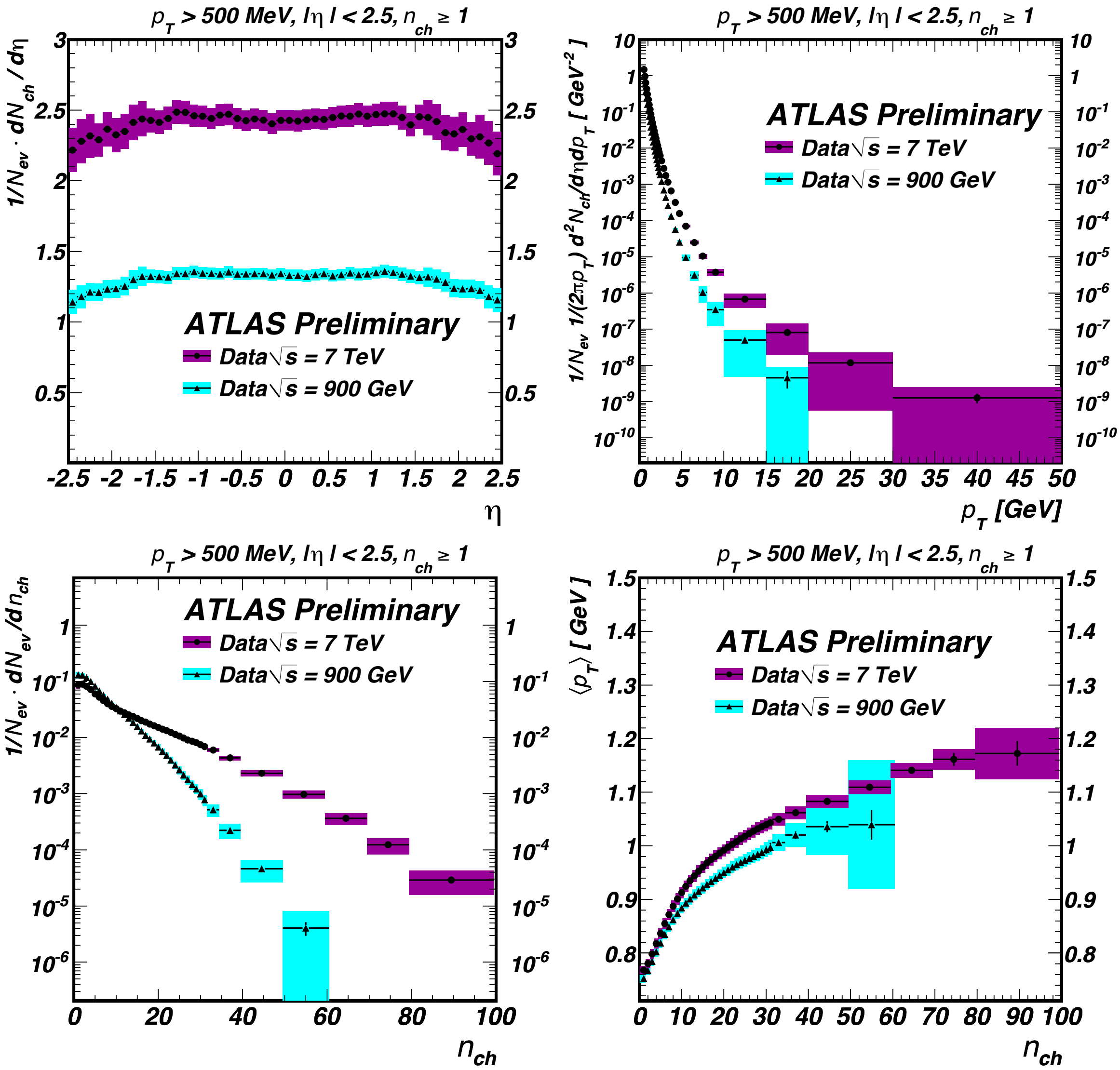



Figure 10.7: Charged particle multiplicities for events with $n_{ch} \geq 1$ within the kinematic range $p_T > 500$ MeV and $|\eta| < 2.5$ at $\sqrt{s} = 900\ GeV$ compared to the results at $\sqrt{s} = 7\ TeV$. The panels compare the charged particle multiplicities as a function of pseudorapidity (top left) and of the transverse momentum (top right), the charged particle multiplicity (bottom left), and the average transverse momentum as a function of the number of charged particles in the event (bottom right). The triangle and circle dots represent the data at $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$, respectively. The vertical bars represent the statistical uncertainties, while the shaded areas show statistical and systematic uncertainties added in quadrature.

The average charged particle multiplicity extrapolated to $\eta = 0$ is shown as a function of the centre-of-mass energy in Figure 10.8.

The energy dependence of the multiplicity is described within 5% by the ATLAS MC09 tune of PYTHIA. This distribution is highly sensitive to the $p_T^{min}$ cut-off of multiple interactions in the PYTHIA model. In that respect it is interesting to note that the PYTHIA tunes shown here vary significantly in this distribution even though they use very similar values for this cut-off.

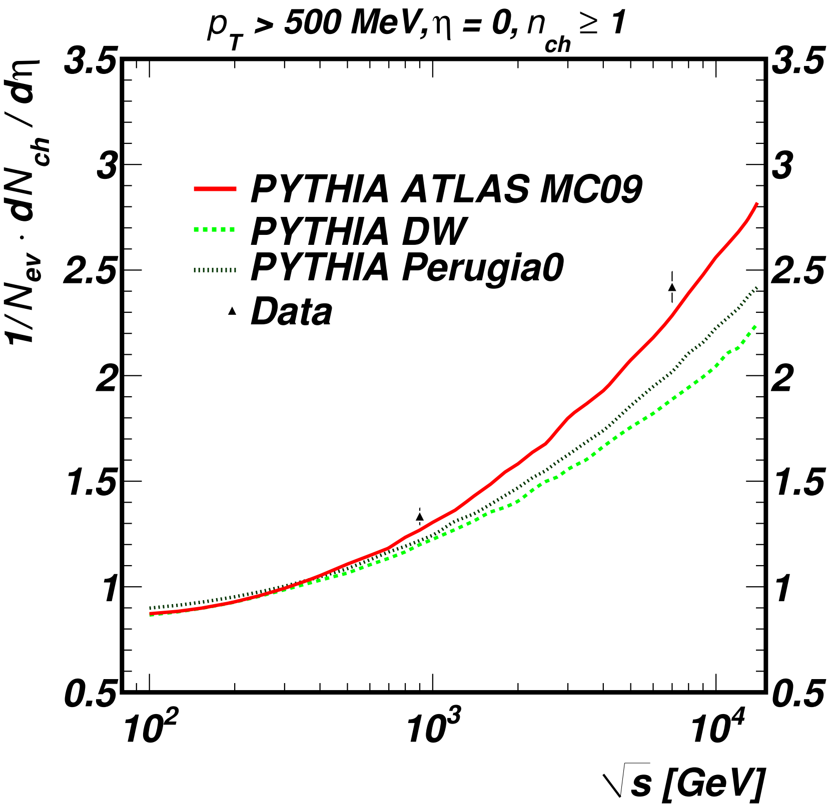



Figure 10.8: The average charged particle multiplicity per unit of rapidity at $\eta = 0$ for events with $n_{ch} \geq 1$ within the kinematic range $p_T > 500$ MeV as a function of the centre-of-mass energy. The triangles represent the data. Predictions from popular PYTHIA tunes are also shown.

## 10.2 Comparison with results from other experiments

The ATLAS minimum bias data at $\sqrt{s} = 900$ GeV can be compared to measurements from other experiments. These results must be, however, converted to the same kinematic region as in the ATLAS minimum bias measurements.

Charged particle multiplicity as a function of the transverse momentum ($p_T$) at 900 GeV in the kinematic range $p_T > 500$ MeV and $|\eta| < 2.5$ is shown in Figure 10.9 compared with the results from the other experiments, which were not adopted to the phase space of the ATLAS measurement on this plot.

The CMS [30] results at the same centre-of-mass energy are superimposed. The number of charged particles in the CMS data is consistently lower than the data presented in this analysis. This offset is expected from the CMS measurement definition of NSD events, where events with $n_{ch} = 0$ enter the normalisation and the number of lower transverse momentum particles are reduced by the subtraction of the PYTHIA single diffractive component.

The UA1 [5] results, normalised by their associated cross-section measurement, are also overlaid. They are approximately 20% higher than the present LHC data. A shift in this direction is expected from the double-arm scintillator trigger requirement used to collect the UA1 data, which rejected events with low charged particle multiplicities.

To compare more directly the present data with results from CMS, the mean charged particle density is calculated in the range $|\eta| < 2.4$ and a model dependent correction is applied to form an NSD particle density. For the calculation of the NSD value the PYTHIA DW tune is selected due to its similarity with the tune used in the CMS analysis. This generator set-up is used to produce a correction for the removal of the fraction of single diffractive events. The resulting value is consistent with the CMS measurement in the kinematic range of $p_T > 500$ MeV and $|\eta| < 2.4$.

Once ATLAS measurements are redefined in terms of the CMS results the two agree within the reported systematic errors.

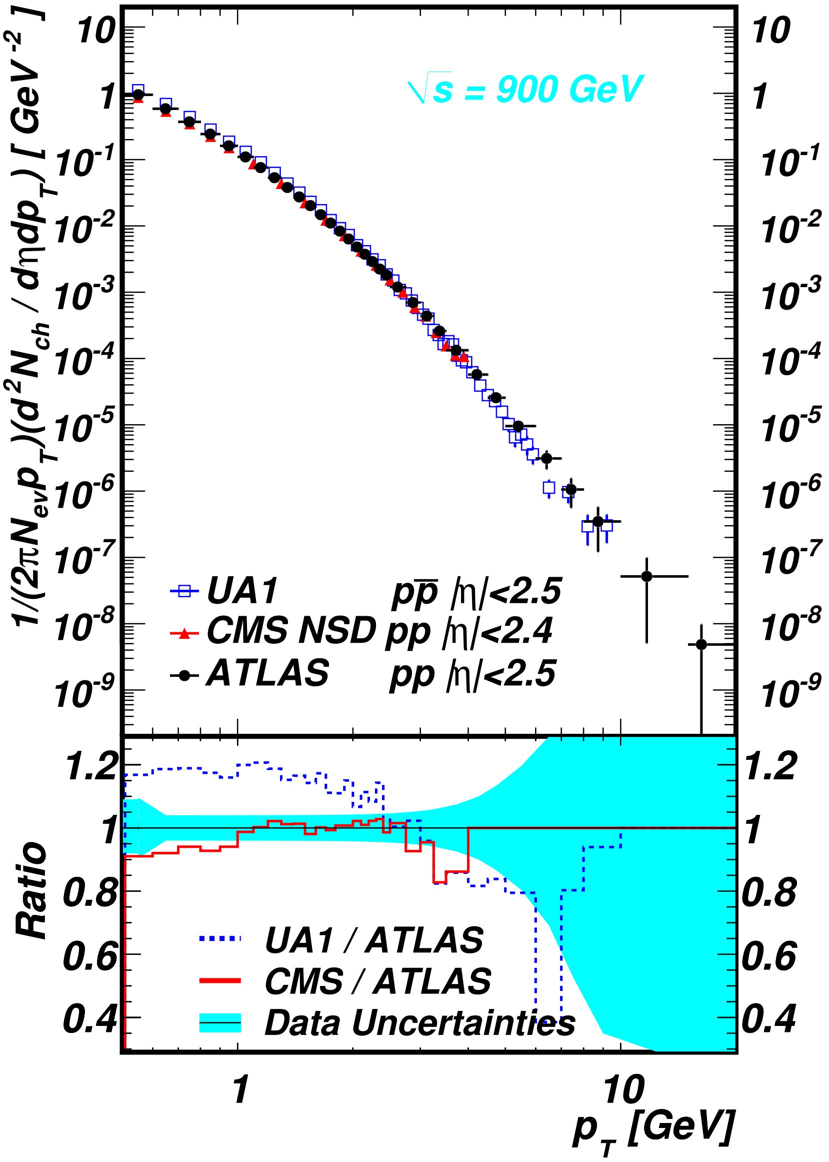



Figure 10.9: The measured $p_T$ spectrum of charged particle multiplicities. The ATLAS pp data (black dots) are compared to the UA1 $p\bar{p}$ data (blue open squares) and CMS NSD pp data (red triangles) at the same centre-of-mass energy.

# Chapter 11

# Conclusion

Inclusive charged particle distributions have been measured in pp and proton antiproton collisions at a range of different centre-of-mass energies in the past. Many of these measurements have been used to constrain phenomenological models of soft hadronic interactions and to predict properties at higher centre-of-mass energies.

Most of the previous charged particle multiplicity measurements are obtained by selecting data with a double-arm coincidence trigger, thus removing large fractions of diffractive events. The data are then further corrected to remove the remaining single-diffractive component. This selection is referred to as NSD. In some cases, designated as inelastic non-diffractive, the residual double-diffractive component is also subtracted. The selection of NSD or inelastic non-diffractive charged particle spectra involves model-dependent corrections for the diffractive components and for effects of the trigger selection on events with no charged particles within the acceptance of the detector.

The measurement presented in this thesis implements different strategy, which uses a single-arm trigger overlapping with the acceptance of the tracking volume. Results are presented as inclusive inelastic distributions, with minimal model-dependence, by requiring one charged particle within the acceptance of the measurement.

The measurements from pp collisions at $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$ recorded with the ATLAS detector at the LHC are performed in this thesis. Data were collected in December 2009 and in March 2010 using dedicated minimum bias trigger. The charged particle multiplicity, its dependence on transverse momentum and pseudo-rapidity, and the relationship between mean transverse momentum and charged particle multiplicity are measured for events with at least one charged particle in the kinematic range $|\eta| < 2.5$ and $p_T > 500\ MeV$. The procedures used in both analysis at $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$ are very similar.

The measurements are compared to MC models of pp collisions and to results from other experiments at the same centre-of-mass energy. The charged particle multiplicity per event and unit of pseudorapidity at $\eta = 0$ is measured to be $1.333 \pm 0.003$ (stat.) $\pm 0.027$ (syst.) at $\sqrt{s} = 900\ GeV$ and $2.418 \pm 0.004$ (stat.) $\pm 0.076$ (syst.) at $\sqrt{s} = 7\ TeV$.

In both scenarios, at $\sqrt{s} = 900\ GeV$ and $\sqrt{s} = 7\ TeV$, the predictions by various MC models differ by 20% and all underestimate the particle production. The charged particle pseudo-rapidity density distribution is best described by the ATLAS MC09c tune which has the same shape as data and is about 5% lower. None of the models are able to describe the data precisely over the whole $p_T$ spectrum. An expected tuning of soft-QCD MC models will follow.

# Acknowledgements

From experience I can tell you that these last pages of a PhD thesis are the most widely read pages of the entire publication. It is here where you think that you will find out whether you have meant something in the life of the PhD candidate. While this may be true to some level, you will have to weigh my verdict with the disturbingly low level of sanity left in this PhD candidate after years of studying.

But for now you will not know what goes through my mind.
Just a simple, thanks mom.

# Bibliography


[1] B. Oliver Sim, C Paul, P. Lebrun, S. Myers, R. Ostojic, J. Poole, P. Proudlock, LHC Design Report. Geneva : CERN, 2004.

[2] G Walter, S Andreas (1994). Quantum Chromodynamics. Springer. ISBN 0-387-57103-5.

[3] S. Abachi et al. D0 Collaboration, Observation of the Top Quark Phys. Rev. Lett. 74, 26322637 (1995).

[4] J.D. Bjorken, S.D. Drell, Relativistic Quantum Mechanics, 1964.

[5] C. Albajar, et al., UA1 Collaboration, A Study of the General Characteristics of $p\bar{p}$ Collisions at $\sqrt{s} = 0.2$ TeV to 0.9 TeV, Nucl. Phys. B335 (1990) 261.

[6] T. Aaltonen, et al., CDF Collaboration, Measurement of Particle Production and Inclusive Differential Cross-Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV, Phys. Rev. D79 (2009) 112005.

[7] P. Skands, D. Wicke, Non-perturbative QCD Effects and the Top Mass at the Tevatron, Eur. Phys. J. C52 (2007) 133.

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

[9] G. Aad, et al., ATLAS Collaboration, ATLAS Monte Carlo Tunes for MC09 (2010). ATL-PHYS-PUB-2010-002.

[10] P. Z. Skands, The Perugia Tunes, 2009. ArXiv:0905.3418v1 [hep-ph].

[11] M. G. Albrow, et al., Tevatron-for-LHC Report of the QCD Working Group, 2006. ArXiv:0610012 [hep-ph].

[12] R. Engel, Photoproduction within the two Component Dual Parton Model. 1. Amplitudes and Cross-Sections, Z. Phys. C66 (1995) 203-214.

[13] Z. Marshall, The ATLAS Simulation Software, 2008. ATL-SOFT-PROC-2008-001.

[14] S. Agostinelli, et al., GEANT4 Collaboration, GEANT4: A Simulation toolkit, Nucl. Instr. Meth. A506 (2003) 250-303.

[15] ATLAS Collaboration, The ATLAS Experiment at the CERN Large Hadron Collider, JINST 3 (2008) S08003.

[16] By ATLAS Collaboration, Tile Calorimeter Technical Design Report, CERN/ LHCC/96-42, 1996.

[17] 2006; A PMT-Block Test Bench. arXiv:phisics/0605074 v1.

[18] By ATLAS Collaboration, Charged Particle Multiplicities in pp Interactions at $\sqrt{s}$ =900 GeV Measured with the ATLAS Detector at the LHC. Phys.Lett.B688:21-42,2010.

[19] By ATLAS Collaboration, ATLAS Detector and Physics Performance Technical Design Report, CERN/LHCC/99-14,15 (1999).

[20] A. Capella et al., DPM1, Phys. Lett. 81B (1979) 68.

[21] A. Capella et al., DPM2, Phys. Rep. 236 (1994) 225.

[22] T. Sjostrand and M. van Zijl Phys. Rev. D 36 (1987) 2019.

[23] A. Moraes, C. Buttar, and I. Dawson, Prediction for Minimum Bias and the Underlying event at LHC Energies, Eur. Phys. J. C50 (2007) 435466.

[24] A. Sherstnev and R. S. Thorne, Parton Distributions for LO Generators, Eur. Phys. J. C55 (2008) 553575, arXiv:0711.2473 [hep-ph].

[25] J. Pumplin et al., New Generation of Parton Distributions with Uncertainties from Global QCD Analysis, JHEP 07 (2002) 012, arXiv:hep-ph/0201195.

[26] R. Field, CDF Run-2 Monte Carlo Tunes, Tech. Rep. CDF/PHYS/JET/ PUBLIC/8547.

[27] By ATLAS Collaboration, Charged Particle Multiplicities in pp Interactions at $\sqrt{s}$ =7 TeV Measured with the ATLAS Detector at the LHC. ATLAS-CONF-2010-024.

[28] By ATLAS Collaboration, ATLAS computing : Technical Design Report. ATLAS-TDR-017 ; CERN-LHCC-2005-022.

[29] By ATLAS Collaboration, Luminosity Determination Using the ATLAS Detector. ATL-ATLAS-CONF-2010-060.

[30] V. Khachatryan, et al., CMS Collaboration, Transverse Momentum and Pseudorapidity Distributions of Charged Hadrons in pp Collisions at $\sqrt{s} = 0.9$ and 2.36 TeV, JHEP 02 (2010) 041.