UNIVERSITÀ DEGLI STUDI DI PISA

Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di Laurea in Scienze Fisiche
Curriculum Fisica delle Interazioni Fondamentali

TESI DI LAUREA SPECIALISTICA

Study of inelastic processes in proton-proton collisions at the LHC with the TOTEM Experiment

Relatore: Prof. S. Lami
Candidato: P. Brogi
Correlatore: Prof. G. Latino

Anno Accademico 2010/2011
Sessione Autunnale - 20 Settembre 2011
# Contents

## Introduction

## TOTEM at the Large Hadron Collider

1.1 The machine ................................. 7

1.2 The TOTEM detectors .......................... 10

1.2.1 The Roman Pots ............................ 11

1.2.2 The T1 telescope ........................... 13

1.2.3 The T2 telescope ........................... 15

1.3 TOTEM offline software and T2 reconstruction chain .......... 18

## The TOTEM physics programme

2.1 The total proton-proton Cross-Section ......................... 25

2.2 The elastic cross-section measurement ....................... 29

2.3 The diffractive process .......................... 31

## Tuning of Geant4 simulation

3.1 A brief introduction to Geant4 .......................... 35

3.2 Geant4 cuts tuning ................................ 37

3.2.1 Geant parameters tuning ....................... 39

3.2.2 Getting the correct list of selected volumes ............ 42

3.2.3 Geant parameter optimization in terms of CPU time .... 45

3.3 Optimization of the simulated geometry in the forward region .... 49

3.3.1 Geometry mismatches and simulation improvements .......... 49

3.4 Comparison between data and the improved simulation .......... 60
3.4.1 Conclusions and possible explanation ........................................... 63

4 Noise contribution in T2 detector ................................................................. 67
   4.1 The noise characterization ................................................................. 68

5 A first study of inelastic events with the T2 detector ................................ 81
   5.1 Study of Single Diffractive events ..................................................... 83
   5.2 Study on Double Diffractive events ................................................... 92
   5.3 A brief looking upon Non Dissociative events ..................................... 99

Conclusions ................................................................................................. 103

Acknowledgements ..................................................................................... 105

Bibliography ............................................................................................... 107
Introduction

The TOTEM [1] experiment, located into the CMS cavern at the CERN Large Hadron Collider (LHC), is one of the six experiments that are investigating high energy physics at this new machine. In particular TOTEM has been designed for Total cross-section, Elastic scattering and diffraction dissociation Measurements. The total proton-proton cross-section will be measured with the luminosity-independent method based on the Optical Theorem. This method will allow a precision of 1÷2% at the center of mass energy of 14 TeV. In order to reach such a small error it is necessary to study the p-p elastic scattering cross-section \( \frac{d \sigma}{dt} \) down to \(|t| \sim 10^{-3} \text{ GeV}^2\) (to evaluate at best the extrapolation to \(t = 0\)) and, at the same time, to measure the total inelastic interaction rate. For this aim, elastically scattered protons must be detected at very small angles with respect to the beam while having the largest possible \(\eta\) coverage for particle detection in order to reduce losses of inelastic events.

In addition, TOTEM will also perform studies on elastic scattering with large momentum transfer and a comprehensive physics programme on diffractive processes (partly in cooperation with CMS), in order to have a deeper understanding of the proton structure.

For these purposes TOTEM consists in three different sub-detectors: two gas based telescopes (T1 and T2) for the detection of inelastic processes with a coverage

\[ s = (p_1+p_2)^2 = (p_3+p_4)^2 \]
\[ t = (p_1-p_3)^2 = (p_2-p_4)^2 \]
\[ u = (p_1-p_4)^2 = (p_2-p_3)^2 \]

Therefore \(s\) represents the square of the c.m. energy, while \(t\) is the four-momentum transfer squared.

\[ \eta = -\ln(\tan \frac{\theta}{2}) \]

where \(\theta\) is the polar angle of the scattered particle with respect to the beam direction.

---

1In a two body scattering \(a + b \rightarrow a + b\), defining the four-momentums of ingoing \((p_1, p_2)\) and outgoing \((p_3, p_4)\) particles, the kinematics can be described using the Lorentz invariant Mandelstam Variables \((s, t, u)\), that are defined as:

2The pseudorapidity is defined as \(\eta = -\ln(\tan \frac{\theta}{2})\), where \(\theta\) is the polar angle of the scattered particle with respect to the beam direction.
in the range of $3.1 \leq |\eta| \leq 6.5$ on both sides of the interaction point 5 (IP5), and silicon based detectors for the elastically scattered protons, located in special movable beampipe insertions called Roman Pots (RPs), at about 147 m and 220 m from the interaction point.

The work done by the candidate reported in this thesis mainly consists in three subjects: the tuning of the simulation for the T2 inelastic telescope, the study of the noise of the T2 detector and a preliminary study concerning the detection performance for inelastic events. In the following, the first chapter describes the TOTEM experiment and the LHC machine, with a particular attention to the T2 telescope and its analysis software, being of critical importance for the work of this thesis. The second chapter introduces the physics programme of the TOTEM experiment. Chapter three describes the tuning of Geant4 parameters and the improvement of the simulated geometry for the T2 detector, while chapter four summarizes an important and demanding study on the detector noise. Finally in chapter five some preliminary studies on inelastic processes are presented, in order to show the perspective for the TOTEM experiment to perform the measurement of the inelastic cross section in a wide kinematic range.
Chapter 1

TOTEM at the Large Hadron Collider

The TOTEM experimental apparatus, consisting in three different sub-detectors, is located at the Interaction Point 5 (IP5) of the CERN (European Organization for Nuclear Research) Large Hadron Collider (LHC), sharing it with the CMS experiment. In this chapter the TOTEM detectors will be described, after a brief overview of the machine. Being the T2 inelastic telescope the main subject of this thesis work, more emphasis will be dedicated to this detector in the following.

1.1 The machine

The LHC, originally started up in September 2008, is the biggest and most powerful particle collider actually operating. It is a circular accelerator of about 27 Km of circumference, located underground (50 to 175 m) into the tunnel of its precursor LEP. It was designed in order to collide two counter rotating beams of protons or heavy ions. For proton-proton collisions it is foreseen to reach a peak luminosity up to $10^{34}$ cm$^{-2}$s$^{-1}$ at a center of mass (C.M.) energy of 14 TeV. It is currently running up to $10^{32}$ cm$^{-2}$s$^{-1}$ and at a 7 TeV C.M. energy. While the design energy is planned to be reached in 2014.

In reality LHC is the final stage of an accelerator complex (figure 1.1) located in the north-west suburbs of Geneva on the French-Swiss border. The first step of the
chain is provided by the LINAC2 linear accelerator, where protons obtained from the dissociation of hydrogen are accelerated up to 50 MeV. They are then sent into the PS Booster that brings them up to 1.4 GeV. Then, inside the Proton Synchrotron (PS) they reach an energy of 26 GeV and are ready for the last pre-acceleration stage in the Super Proton Synchrotron (SPS) which accelerates protons up to 450 GeV before injecting them into the LHC ring. Where they are brought to the final energy (presently 3.5 TeV per beam), before colliding in the four provided collision points, where are located the six CERN experiments (ALICE, ATLAS/LHCf, CMS/TOTEM, LHCb) with their detectors.

In order to keep two counter-rotating proton beams, the machine needs two separated rings with opposite magnetic fields to bend same charge particles rotating on opposite directions. Moreover, in order to bend 7 TeV protons a magnetic field of 8.36 Tesla is required. The goal is achieved using two different superconducting dipoles housed in the same yoke, cooled down to 1.9 K with superfluid helium. The whole accelerator is composed of 1296 superconducting dipoles (bending magnets)
and more than 2500 other magnets used to guide and focus the beams around the ring. The main parameters of the machine are reported in table 1.1 with their nominal values [2].

<table>
<thead>
<tr>
<th>Data</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum proton energy</td>
<td>[GeV]</td>
<td>7000</td>
</tr>
<tr>
<td>Number of particles per bunch</td>
<td></td>
<td>1.15 \cdot 10^{11}</td>
</tr>
<tr>
<td>Number of bunches</td>
<td></td>
<td>2808</td>
</tr>
<tr>
<td>Circulating beam current</td>
<td>[A]</td>
<td>0.582</td>
</tr>
<tr>
<td>Peak luminosity in IP5</td>
<td>\text{cm}^{-2}\text{s}^{-1}</td>
<td>1.0 \cdot 10^{34}</td>
</tr>
<tr>
<td>Inelastic cross section</td>
<td>[mb]</td>
<td>60.0</td>
</tr>
<tr>
<td>Total cross section</td>
<td>[mb]</td>
<td>100.0</td>
</tr>
<tr>
<td>Beam current lifetime</td>
<td>[h]</td>
<td>18.4</td>
</tr>
<tr>
<td>Ring circumference</td>
<td>[m]</td>
<td>26658.883</td>
</tr>
<tr>
<td>Revolution frequency</td>
<td>[kHz]</td>
<td>11.245</td>
</tr>
<tr>
<td>RF frequency</td>
<td>[MHz]</td>
<td>400.8</td>
</tr>
<tr>
<td>Synchrotron frequency (during collision)</td>
<td>[Hz]</td>
<td>21.4</td>
</tr>
<tr>
<td>Half crossing angle at IP1 and IP5</td>
<td>\text{µrad}</td>
<td>±142.5</td>
</tr>
<tr>
<td>(\beta) at IP1 and IP5 (during collision)</td>
<td>[m]</td>
<td>0.55</td>
</tr>
</tbody>
</table>

Table 1.1: Main parameters of the LHC machine at nominal c.m. energy (14 TeV).

One of the most important parameters shown here is the luminosity \(\mathcal{L}\), representing the factor of proportionality between the event rate \(R\) and the interaction cross-section \(\sigma\):

\[
R = \mathcal{L}\sigma
\]

So it is easy to understand why, in order to observe phenomena with a very low cross-section, it is important to reach the highest possible luminosity, which can be defined as [3]:

\[
\mathcal{L} = fn_b \frac{N_1 N_2}{4\pi\sigma_x \sigma_y}
\]

In this equation \(\sigma_x\) and \(\sigma_y\) represent the transverse gaussian beam profiles at the IP in the horizontal and vertical directions. \(N_1, N_2\) represent the number of protons in the colliding bunches, \(f\) is the frequency of revolution of bunches and \(n_b\) the number of bunches. The equation is not exact for calculating \(\mathcal{L}\) at LHC but it means that a higher luminosity can be reached with small transverse size bunches at IP or a high number of bunches (\(f\) depends only on the accelerator length) or highly populated bunches. However anyone of these requirements can cause several
problems. Higher focussed bunches lead to severe “beam-beam effects” (when two bunches cross, the particles are deflected by the strong electromagnetic field, this deflection is stronger for denser bunches, and can lead to particle losses). Increasing the number of particles in each bunch results in more event pile-up and this is to avoid for a better understanding of the physics process. Furthermore, the bunch crossing rate is limited by the time resolution of the detectors and read-out systems employed.

1.2 The TOTEM detectors

The TOTEM experiment is composed by three different detectors: the two telescopes T1 and T2, based on CSC (Cathode Strip Chamber) and GEM (Gas Electron Multiplier) technology, respectively; and the Roman Pots (RPs) equipped with silicon detectors. The three detectors are located (see figure 1.2) symmetrically on both sides of the interaction point IP5, the same shared with the CMS experiment. The telescopes are located at 9 m and 13 m from the interaction point, while the RPs are located in special vacuum insertions along the beam-pipe at 147 m and 220 m from IP5. The detectors designed for a particular purpose have a specific acceptance region; in particular the TOTEM physics programme requires a good acceptance for angles very close to the beam axis. The pseudo-rapidity coverage for T1 and T2 is $3.1 < |\eta| < 6.5$, and the RPs placed inside the vacuum pipe allow the detection of elastically scattered very close to the beam (till few $\mu$rad). The data acquisition system is designed to be compatible with CMS to have the possibility of a common data taking in order to combine TOTEM and CMS, therefore obtaining the largest acceptance (in eta) detector ever built. The two inelastic telescopes have a $2\pi$ coverage in $\phi$ and a good efficiency in order to minimize losses of non-diffractive and minimum bias events. They are designed to ensure the detection of about the 95% of all inelastic events having charged particles within their geometrical acceptance (about 99.5% of all non-diffractive events and 84% of all diffractive processes). Even if the telescopes are outside the central region of the CMS magnetic field and cannot provide information about the momentum of tracked particles, they are in front of
two CMS calorimeters, HF for T1 and Castor for T2, respectively. Therefore the combination of this two kind of detectors could permit a more complete study of the diffractive processes, low-x phenomena and particle/energy flows in the very forward region. The read-out task of all the TOTEM detectors is provided by the VFAT2 (Very Forward ATLAS and TOTEM chip) [4], a front-end ASIC (Application Specific Integrated Circuit) designed in CMOS technology for the TOTEM experiment itself to process the signals and marked by trigger capability.

### 1.2.1 The Roman Pots

Silicon detectors are placed inside each secondary vacuum insert, called “pot”. These special pots are moved into the primary vacuum through a bellow. This device allows to physically separate the detectors from the primary vacuum, in order to preserve it from an uncontrolled out-gassing of the materials. This experimental technique is well known since it was introduced at the ISR and it has been successfully employed
1.2 The TOTEM detectors

Figure 1.3: Schematic view of a Roman Pot station

in other colliders like SppS, TEVATRON, RIHC and DESY. Moreover the use of movable inserts is useful because it allows to retract the detectors in a safe position when the beam is in an unstable condition, avoiding useless risk and exposure to radiation for the silicon detectors. There are two RP stations, like the one depicted in figure 1.3, for each side of IP5, placed at a distance of about 147 m and 220 m along the beam-pipe, symmetrically on both sides. A magnetic dipole between the two RP stations provides a magnetic spectrometer which helps proton momentum reconstruction. Each RP station is composed by two RP units (figure 1.4) separated by a distance that allows local track reconstruction and trigger selection by the track angle. A RP unit consists in 3 pots, two approaching the beam vertically and one horizontally, the scheme of silicon detectors displacement is shown in figure 1.5. This configuration was chosen to provide the best reconstruction of the fractional momentum loss ($\xi$) of diffractive scattered protons. Looking inside the pots at the detectors itself, these are constituted by a stack of 10 planes of silicon “edgeless” devices. These are single-sided AC $p^+\cdot n$ microstrip detectors 300 $\mu$m thick with 512 strips and a pitch of 66 $\mu$m. Half of the silicon devices have their strip oriented at an angle of $+45^\circ$, with respect to the edge facing the beam, and the others to $-45^\circ$. This structure allows a single hit resolution of about 20 $\mu$m. The special
1.2 The TOTEM detectors

1.2.2 The T1 telescope

The T1 telescope has two arms, one for each side of IP5, and it is installed into the CMS End Caps at a distance of 7.5 to 10.5 m. Each arm surrounds the beam-pipe and has a coverage in pseudo-rapidity of $3.2 < |\eta| < 4.7$. The detector is based on Cathode Strip Chamber (CSC) technology, the CSC being a multi-wire proportional chamber, with a read-out made by a segmented cathode. Each plane of T1 consists in six trapezoidal CSCs. Five of these planes, equally spaced in $z$, build an arm of the telescope. In order to have a better pattern recognition for track reconstruction and to reduce the material concentration in front of the CMS calorimeter (Hadron Forward) the six trapezoidal CSCs of each plane are tilted with respect to each other.
1.2 The TOTEM detectors

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full gas gap</td>
<td>10 mm</td>
</tr>
<tr>
<td>Wire spacing</td>
<td>3 mm</td>
</tr>
<tr>
<td>Wire diameter</td>
<td>30 $\mu$m</td>
</tr>
<tr>
<td>Strip pitch</td>
<td>5 mm</td>
</tr>
<tr>
<td>Strip width</td>
<td>4.5 mm</td>
</tr>
<tr>
<td>Chamber thickness</td>
<td>43 mm</td>
</tr>
</tbody>
</table>

Table 1.2: Basic parameters of T1 Cathode Strip Chambers

by a small angle varying from $-6^\circ$ to $+6^\circ$ in steps of $3^\circ$. The read-out boards on both sides of a chamber are segmented in strips and are rotated to $+60^\circ$ and $-60^\circ$ with respect to the anode wires. That allows (according to beam test studies) a spatial resolution of about 0.8 mm, when using a digital read-out. Moreover the gas mixture employed is Ar/CO$_2$/CF$_4$ in a ratio 40%/50%/10%, and with this kind of mixture and a gas gap of 10.0 mm the time response for this detector (even if inherently slow) is compatible with the rates required by TOTEM. Aging studies for this detector have shown no loss of performance after an irradiation equivalent to a total charge integrated on the anode wire of 0.065 C/cm, which corresponds to 5 years of running at a luminosity of $10^{30}$ cm$^{-2}$s$^{-1}$. In table 1.2 are summarized some important parameters of the T1 Cathode Strip Chambers, and in figure 1.6 are shown an arm of T1 (left) and a schematic view of one chamber of the telescope (right).

Figure 1.6: Left: a T1 telescope arm. Right: schematic view of anodic wires and cathodic strips displacement.
1.2 The TOTEM detectors

1.2.3 The T2 telescope

The T2 telescope is composed by 40 planes of Gas Electron Multiplier (GEM) detectors, with an angular coverage of 192° each. The schematics of one of this plane is shown in figure 1.7; it displays the detailed shape that allows the detector to enclose the beam-pipe at its center. Moreover, cooling and read-out sectors are also visible in the picture. One quarter of the telescope is made by 10 planes aligned and assembled back-to-back in five pairs, each distant 91.0 mm from the other, for a total length of the quarter of about 40 cm (supports excluded). Two quarters form an arm of the telescope with a coverage of 360° in φ and an overlap region of 12°, to minimize the edge inefficiency. The two arms are placed on each side of IP5 located at ±13.5 m inside the shielding behind HF and before the Castor calorimeter. More precisely, the Z position of the first GEM plane with respect to the IP is 13.83 m. From figure 1.8 it is possible to evaluate the T2 position with respect to the CMS calorimeters and the ion pump station placed in the beam-pipe, just in front of the T2 telescope. Moreover in figure 1.9 a 3D schematic view of one arm of the T2 detector is shown. The T2 coverage in pseudo-rapidity is $5.3 \leq |\eta| \leq 6.5$, the resolution in $\eta$ is good, down to 0.04, and it allows a good capability in discriminat-
1.2 The TOTEM detectors

Figure 1.8: Location of one arm of the T2 detector inside the shielding behind HF, in front of the Castor calorimeter.

Figure 1.9: 3D view of one arm of the T2 detector.
The TOTEM detectors

ing against beam-gas background and secondary particles produced in interactions with the beam pipe. The GEM technology used for the T2 telescope ensures a high rate capability, good spatial resolution and good radiation hardness. This kind of detectors, invented about a decade ago by Fabio Sauli [6], are characterized by a very high efficiency in detecting charged particles and they are used also in other CERN experiments like COMPASS and LHCb.

The structure of the GEM chambers is based on the “triple-GEM” scheme adapted also in COMPASS, in which three GEM foils are assembled in cascade, as shown in figure 1.10, where the transversal view displays the composition of a semicircular detector, made by a stack of three GEM foils, separated by 2 mm insulator spacers and mounted on the honeycomb supports. This configuration is useful in order to obtain an high gain reducing the discharge probability (below $10^{-12}$). A GEM foil is composed by a 50 $\mu$m polyimide sheet coated with 5 $\mu$m copper on both sides. On the foil there is a high density of holes, obtained by a photo-lithographic method, with a double conical shape (the distance between the holes is 140 $\mu$m). The diameter of the holes is 65 $\mu$m in the middle of the GEM foil and 80 $\mu$m at the surface. A charged particle crossing the chamber ionizes the gas that fills the drift volume (a mixture of Ar/CO$_2$ at 70%/30%) producing primary electrons, which are carried by an electric field of about 2.4 KV/cm towards the holes of the top GEM-plane, where an electric field of about 50 KV/cm, which generates the electron multiplication, is present. This field is achieved by applying a voltage (about 400V) between the two copper layers of a foil. For this configuration the factor of multiplication is about

![Figure 1.10: Internal structure of a triple-GEM detector.](image1)

![Figure 1.11: Read-out board of a T2 triple-GEM detector.](image2)
20 and the electrons produced inside the channels were driven by a field of about 3.6 KV/cm through the transfer zone to the next GEM planes (where the following electron multiplications happen) till finally the charge is collected on the read-out board. This board was specifically designed for TOTEM, and it has an inner radius of 42.46 mm and an outer radius of 144.46 mm. The structure as shown in figure 1.11 is composed by two layers of 15 µm copper, separated by a polyimide foil of 50 µm. The two layers have different patterns, one is divided in strips while the other in pads. The first is divided in 2 sectors of 256 concentric circular strips, 80 µm wide and with a pitch of 400 µm. Each sector covers an azimuthal angle of 96° and the strip segmentation allows track radial reconstruction. The second layer instead is segmented in pads, which provide level-1 trigger information and track azimuthal angle reconstruction. The pads form a matrix of 24x65 elements, varying in size from 2 x 2 mm² to 7 x 7 mm², in order to have a constant Δη x Δφ ~ 0.06 x 0.05 rad. Beam tests on detectors have shown a spatial resolution in radial coordinates of about 100 µm (with digital VFAT read-out), while the time resolution achievable with the electric field reported above is about 18/20 ns. Concerning the detector aging, tests on COMPASS triple-GEMs have shown that a charge up to 20 mC/mm² can be integrated on the read-out board without major effects. This corresponds to running for at least 1 year at luminosities of 10^{33} cm^{-2}s^{-1}. All these features make the triple-GEM technology a proper choice for the T2 telescope requirements.

1.3 TOTEM offline software and T2 reconstruction chain

Since the use of the TOTEM offline software is mandatory for the thesis work, in this section a brief description of this tool is provided. Particular attention has been paid to the T2 reconstruction chain, since it is fundamental, to understand this thesis work, to know how the particle induced signals (real or simulated ones) collected by the detector are treated at the analysis level to reconstruct useful observables like: clusters, hits and tracks.
Looking at the software structure, TOTEM is using a C++ based framework developed by the CMS experiment (CMSSW) [7]. This allows to reconstruct and record physics and simulation events ensuring a full compatibility of TOTEM and CMS data processing in future analyses studies. This framework consists of an Event Data Model (EDM), services needed by the simulation and reconstruction modules that process event data. The EDM is based around the concept of an Event. This is a C++ object container for all the information coming from real data acquisition or from physical process simulation. The Event starts as a collection of raw data (signals) from detectors or as a collection of the generated particles in a Monte Carlo (MC) simulated event. Then during the processing (via reconstruction modules) the Event is used to pass the data from one module to the next, to access them and to store the products of processing in objects. All these objects contained in the Event may be stored (collectively or individually) in ROOT format files, and are thus directly readable in ROOT [8].

Effectively once that one has a data file, from a real detector acquisition or the simulation of a physical process, the interesting observables can be reconstructed on it and then an analysis on these observables can be performed. The reconstruction is identical for both real data and simulated processes, and is performed for the T2 detector in four main steps: clusterization, hit reconstruction, road finding and tracking [9]. Anyway if one wants to simulate the detector response for a physical process there are three steps to do before the reconstruction; since the simulated event has to be generated, propagated and digitized. Because of the importance of the simulation in our thesis work, we spend few words also in the description of these steps. The event generation is handled by a Monte-Carlo generator: Pythia6 [10], Pythia8 [11] and Phojet [12] are common ones and are all used in this thesis. These generators allow to produce the final state of a proton-proton collision for a wide variety of physical processes (and C.M. energies). Moreover there is also the possibility to use a Particle Gun, that allows to generate single or multiple particles with fixed (or alternatively flat distributed) values of $\eta, \phi$ and energy at the IP. After the generation of the physical process, it has to be propagated from the IP to the detectors, simulating the interaction of the particles with matter as
well as the effect of the magnetic field. This is performed by a software tool named Geant4 [13] (a more comprehensive description on it is reported in section 3.1). Given the Geant4 simulation of particle entry and exit points in the detector active volumes and their energy deposition, the digitization step reproduces the electrical response of the detector itself (for instance of a T2 GEM chamber). For what concerns the T2 triple GEM detectors, the proper module inside CMSSW is able to reproduce the digital output signal of the chambers. So, the outputs of the digitization are the pad/strip digital status (ON/OFF) for each telescope plane. This is the same kind of output given by a real data acquisition and it provides the input for the reconstruction process. The first reconstruction step is the clusterization. Since a particle traversing a detector device typically turns ON more than one read-out channel, it’s important to collect all the neighbouring pad/strips in an unique pad/strip cluster. The clusterization algorithm manages to do that and saves all the cluster information that could be useful for the next steps or for analysis purposes. For what concerns pad clusters, only the active pad that touches each other via a side were considered neighbouring and collected in a cluster, while the pad that touches each other via a corner were reconstructed as two different clusters. The most important information saved for each cluster are: the detector ID to which the cluster belongs; the position of the cluster itself, the cluster type (pad or strip) and the cluster size (number of pads/strips in the cluster). The detector ID is an integer number that permits to identify to which plane and quarter of the T2 detector the cluster belongs.

At this point it could be useful to explain the numeration scheme used to identify the T2 detector components. The T2 quarters are numbered from 0 to 3, and are called H0, H1, H2 and H3. H0 is the plus near quarter of the detector, H1 is the plus far, H2 the minus near and H3 the minus far. Near and far means respectively that the quarter is located in the inner side of the LHC ring or in the outer side. While plus and minus means that the quarter is located respectively on the positive half-line of the Z axis\(^1\) or on the negative one. The planes of each quarters are then

\(^1\)The coordinate system we usually refer to in this thesis work has the origin in IP5, the X axis pointing toward the center of the LHC ring, the Y axis pointing to the ground surface and the Z axis along the beam line.
numbered from 0 to 9 starting from the plane nearest to IP5 that is number 0, from the farthest one that is number 9.

After this little digression on the numbering scheme we restart to describe the reconstruction chain, from the hit reconstruction, that is the next step after the clusterization. In fact most of the times that a ionising particle crosses a detector plane it generates both a strip and a pad cluster. For this reason the hit reconstruction algorithm matches the overlapping pad and strip cluster to form a class 1 hit. While the clusters that don’t match with any others are called class 2 hits and become equally part of the hit collection. Then all the information related to an hit are saved in appropriate objects, the most important ones being: the hit position (and resolution on it), the hit class (1 or 2), the composition of a class 2 hit (strip or pad) and all the information inherited by the clusters that composed the hit itself. In the case of a class 1 hit the position information are reconstructed taking advantage from both pad cluster and strip cluster, because the pad cluster has a better angular resolution and the strip cluster a better radial one. This allows to have a more precise measure of the real particle position.

Achieved this step we want to reconstruct also the track (3D straight line) that the particle follows inside the detector. We do this in two distinct steps; we first search for all the “roads” and then we make a linear fitting on these roads to find the tracks. A road is a collection of hits on different planes that are roughly aligned. A road finding algorithm is used, that acts in a few consecutive levels. This algorithm works quarter by quarter and at first level considers only the pad clusters, because of their better granularity and efficiency. At this level starting from the first detector plane (nearest to IP) for each cluster in this plane it computes a raw track with all the clusters in the second plane. After the algorithm checks for each raw track if there is a cluster in the third plane in a position compatible with the crossing point of the raw track in this plane. Compatible means that the center of the pad cluster is more than $2\sigma$ closer to the crossing-point of the track. If there is a such cluster in the third plane this is associated to the other two to form the road, then the search for clusters continues in the next planes and when a compatible one is found it is associated to an existing road. Otherwise the algorithm searches for clusters
compatible with the raw track for other two planes (i.e. fourth and fifth), and if there aren’t none this possible road is discarded and the search for the next starts with a different raw track. After that for the first plane all the possible roads has been computed, the algorithm moves to the second plane and follows the same procedure but using only the pad cluster still not belonging to any road. And so on, until the last plane. In these further steps, in which the algorithm computes the road starting not from the first plane, it also searches back for clusters to associate to the road in the planes that precede the starting one and not only in the plane that follows it. Once that all these pad cluster collections are computed with this method, the algorithm descends in a successive level and associates to the pad cluster roads all the strip clusters compatible with it. Generating in this way a road of hits (of class 1 or class 2). During this procedure it could happen that from the superimposing of a pad and strip cluster more than one hit has been generated. This allows to produce from a road of pad clusters more than one road of hits. For this reason, after that all the hits in the road are computed, the algorithm searches for the sub-collection of hits that, once fitted, generate the best track. Found this, the sub-collection becomes a new road and the software searches for the other possible combination in the old road that fits with a line, and if there are any they become new roads. Once that all the hit roads are computed with this method, the algorithm searches if there are two roads of different quarters that are overlapping: if they are found, then they are merged in a single one.

Found and recorded all the possible hit roads for the four detectors quarter, the next step in reconstruction is to obtain final particle tracks. This is accomplished by the tracker algorithm, that takes as input the roads previously computed and then performs a fit [14] on the hits belonging to them. In a first moment the fit is done on all the road hits, but then if the $\chi^2$ probability$^2$ is greater than 0.01, the algorithm tries to remove the hit with the worst squared deviation and redo the fit. If the new fit passes the cut of 0.01 the track is saved and the hit definitively discarded, otherwise the algorithm tries to discard another hit. The procedure is repeated until

---

$^2$Given a certain Chi-squared ($\chi^2$) and number of degrees of freedom (ndf), is calculated basing on the incomplete gamma function $P(a,x)$ as $1-P(a,x)$ where $a=$ndf/2 and $x=\chi^2/2$. It denotes the probability that an observed Chi-squared exceeds the value $\chi^2$ by chance, even for a correct model.
there are at least three class 1 hits and a class 2 hit remaining in the road. Then, if the $\chi^2$ probability is still greater than 0.01, the fit is redone with all the previously discarded hits and the track saved anyway. Obviously the tracker algorithm saves in appropriate objects a lot of information concerning the track position and the fit parameters. The most important are: the $\theta$ angle of the track associated versor with respect to the Z axis; the track azimuthal angle $\phi$; the minimum approach distance between the reconstructed 3D track and the Z axis $R_0$; the point along the Z axis in correspondence to the minimum approach distance of the track $Z_0$; the track reduced $\chi^2$ and the Chi-squared probability $\chi^2_{\text{prob}}$. 
1.3 TOTEM offline software and T2 reconstruction chain
Chapter 2

The TOTEM physics programme

The TOTEM apparatus was designed to accomplish the TOTal cross section, Elastic scattering and diffractive dissociation Measurement, and thanks to its coverage for charged particles at high $\eta$, it represents the ideal tool for studying forward phenomena. About the 99.5% of all non diffractive minimum bias events and 84% of all diffractive events are triggerable by the inelastic telescope (T1,T2). This is of great importance to perform the total cross-section measurement with the luminosity-independent method, based on the Optical Theorem. Figure 2.1 shows the TOTEM detectors $\eta - \phi$ coverage, where the good coverage at high pseudo-rapidity values can be appreciated. Furthermore figure 2.2 shows the expected particle multiplicity and energy flow as a function of $\eta$ and it is clearly visible how the TOTEM detectors cover a region with high particle multiplicity and energy flow.

2.1 The total proton-proton Cross-Section

The total p-p cross-section ($\sigma_{\text{tot}}$) reflects the various interactions between the colliding particles and their constituents. Therefore a precise measurement of this quantity allows to distinguish between several existing theoretical models of soft proton interactions. The aim is to achieve a precision of about 1÷2 mb corresponding to $\sim$1% of the expected total cross-section at the LHC energy of $\sqrt{s} = 14$ TeV. Indeed, referring to figure 2.3 in which various existing measurements of $\sigma_{\text{tot}}$ are shown, a cross-section typically from 90 to 130 mb is expected, depending on the
2.1 The total proton-proton Cross-Section

model used for the extrapolation. As mentioned above, the luminosity-independent method used by TOTEM for its measurement of the total cross-section relies on the Optical Theorem, written in equation 2.1, where $F(t=0)$ is the forward elastic scattering amplitude for a squared four-momentum transfer $t = 0$ and $k$ is the momentum in the center of mass.

$$
\sigma_{\text{tot}} = \frac{4\pi}{k} \text{Im}\{F(t = 0)\}
$$ (2.1)

From this theorem it follows the equation 2.2 that relates $\sigma_{\text{tot}}$ with the luminosity $\mathcal{L}$, the nuclear part of the elastic cross-section $\frac{dN_{el}}{dt}$ and the parameter $\rho$ (see 2.3).

$$
\mathcal{L}\sigma_{\text{tot}}^2 = \frac{16\pi}{1 + \rho^2} \left. \frac{dN_{el}}{dt} \right|_{t=0}
$$ (2.2)

$$
\rho = \frac{\text{Re}[F(t = 0)]}{\text{Im}[F(t = 0)]}
$$ (2.3)
Figure 2.3: Fits from the COMPETE collaboration [15] to all available p-p and p-\bar{p} scattering data with statistical (blue solid) and total (green dashed) error bands.
The additional following relation, that involves the inelastic rate $N_{inel}$ and the elastic one $N_{el}$, is also clearly valid:

$$N_{el} + N_{inel} = \mathcal{L} \sigma_{tot}$$  \hspace{1cm} (2.4)

Therefore, the previous 2.2 and 2.4 form a system of two equations which can be solved either for $\sigma_{tot}$ or $\mathcal{L}$ and lead respectively to the equation 2.5 and 2.6.

$$\mathcal{L} = \frac{1 + \rho^2}{16 \pi} \cdot \frac{(N_{el} + N_{inel})^2}{dN_{el}/dt|_{t=0}}$$  \hspace{1cm} (2.5)

$$\sigma_{tot} = \frac{16 \pi \cdot dN_{el}/dt|_{t=0}}{1 + \rho^2 (N_{el} + N_{inel})}$$  \hspace{1cm} (2.6)

Hence the quantities to be measured are the following:

- $N_{inel}$ which will be measured by the inelastic telescopes T1, T2 and consists of diffractive ($\sim 18$ mb) and non diffractive minimum-bias ($\sim 65$ mb) events. For this purpose T1 and T2 also provide the reconstruction of the primary vertex in order to discriminate between the beam-beam events and the background ones (mainly from beam-gas interactions and muons halo).

- $N_{el}$ which will be measured by the Roman Pots.

- $dN_{el}/dt|_{t=0}$ which will be measured down to $-t = 10^{-3}$ GeV$^2$ and then extrapolated to $t=0$.

For this measurement it is important that all TOTEM detectors have trigger capability. Moreover particular beam optics conditions are required. In fact the proton detection by the Roman Pot (RP) stations at very small scattering angles ($\sim$ few $\mu$rad) requires special accelerator optics configurations. In particular a high $\beta^*$ function is needed, because the detection of protons elastically scattered so close to the beam axis requires a small beam angular divergence at the interaction point $\sigma_\theta \sim \sqrt{\frac{1}{\beta^*}}$ (small with respect to the scattering angle itself). In addition, the proton revealed in the RP is required to be reasonably away from the beam envelope ($\sigma_{env}$), typically at least $10 \sigma_{env}$. To perform the total cross section measurement
with the required precision of $1 \div 2\%$ an optical configuration with $\beta^* = 1540$ m is desirable, for which a parallel-to-point-focusing condition is reached for the RP placed at 220 m from IP. This condition is important in order to eliminate the dependence on the transverse position of the proton at the collision point, allowing a more precise measure of the scattering angle. Even if a very high $\beta^*$ (1540 m) run is not foreseen for the early stage of LHC, there will be probably soon a run at an intermediate $\beta^* = 90$ m. This will allow TOTEM to make a measurement of the total cross section with a $\sim 5\%$ uncertainty. For this $\beta^* = 90$ m run the systematic error for the measurement will be dominated by the extrapolation of the nuclear elastic cross section to $t = 0$ ($\sim 4\%$ for $-t$ measured down to $-t \sim 10^{-2}$ GeV$^2$), while for the measurement at $\beta^* = 1540$ m the total inelastic rate will give the main systematic uncertainty. In particular the uncertainty will be determined mainly by trigger losses in Single Diffractive events ($\sim 0.8\%$). This occurs when the invariant mass of the fragmented system is quite low (below 10 GeV/$c^2$) so that the particle pseudo-rapidities is beyond the T2 tracker acceptance [16]. Finally the theoretical uncertainty related to the estimate of the $\rho$ parameter is expected to give a relative uncertainty contribution of less than 1.2% [15].

2.2 The elastic cross-section measurement

In order to distinguish among different models of soft proton interactions, the study of large impact parameter collisions such as elastic scattering processes is of great interest. High energy elastic nucleon scattering is a process for which many precise experimental data, covering a large energy range, are available. The predictions for the differential cross section of elastic proton-proton scattering, for an energy of $\sqrt{s} = 14$ TeV, according to several different models [17], are shown in figure 2.4. Depending on the physics of the interaction involved, we can then distinguish several regions in squared transverse momentum $t$:

- $|t| < 10^{-5}$ GeV$^2$: The elastic scattering is dominated by the exchange of one photon; it is called Coulomb region and it follows the Rutherford formula $\frac{d\sigma}{dt} \sim \frac{1}{t^2}$.
2.2 The elastic cross-section measurement

Figure 2.4: Prediction from some different theoretical models for differential cross-section of elastic scattering. Also the acceptance ranges in $|t|$ for different beam optics are shown.

- $10^{-3}$ GeV$^2 < |t| < 0.4$ GeV$^2$: This is called the hadronic region and its theoretical description relies on the single pomeron exchange model. The differential cross section behaviour is roughly exponential $\frac{d\sigma}{dt} \sim e^{-B|t|}$ but the slope is $t$ dependent $B(t) = \frac{d}{dt} \ln \frac{d\sigma}{dt}$. This shows a small model dependent deviation from the exponential behaviour and it leads to a theoretical uncertainty that contributes to the systematic error of the total cross-section measurement. This region is important (together with the “interference” region $10^{-5}$ GeV$^2 < |t| < 10^{-3}$ GeV$^2$) to evaluate the $\frac{dN_{el}}{dt}|_{t=0}$.

- $|t| > 0.4$ GeV$^2$: This is the region in which the diffractive structure of the proton is visible. In fact we can see (from figure 2.4) the shape with maxima and minima that is typical of diffraction.

- $|t| > 1.5 \div 3$ GeV$^2$: The predicted behaviour for the cross section in this region is proportional to $|t|^{-8}$, it relies on the description of central elastic collisions by perturbative QCD for example in terms of three gluon exchange [18]. This region is also useful to test the validity of different models because of the big
difference in their predictions in this range.

With different running conditions and optics the TOTEM experiment will cover a wide t-range, measuring the elastic scattering from $2 \times 10^{-3}$ GeV$^2$ to about 10 GeV$^2$. This is really important to discriminate among different models (figure 2.4) for the differential cross-section of elastic scattering.

2.3 The diffractive process

The diffractive process comprises Single Diffraction (SD), Double Diffraction (DD), “Double Pomeron Exchange” (DPE, or Central Diffraction) and high order processes (“Multi Pomeron Exchanges”). These processes are expected to give a great contribution (about 50 mb together with the elastic scattering) to the total cross section at LHC. All these processes are characterized by a well defined topology in the pseudorapidity-azimuth ($\eta - \phi$) plane, shown in figure 2.5, that allows to distinguish between the various kind of diffractive events or non diffractive Minimum Bias (MB) events. The diffractive processes can be divided into “soft” and “hard”. The former gives almost the overall contribution to the diffractive cross-section, while the latter can be distinguished for the presence of jets in most of their final states. Moreover the main features of these processes are a large, non-exponentially suppressed, rapidity gap and no exchange of quantum numbers between the colliding protons. A large rapidity gap, usually greater than 2 pseudorapidity units while non-exponentially suppressed, means that the probability of finding the gap in the final state is not a strong function of the gap width. An example for these events can be provided by SD processes at HERA ($e + p \rightarrow e + p + X$), where X represents the fragmented system and has an invariant mass $M_x$. In the approximation of $s$, $M_x \gg 1$GeV$^2$ the rapidity gap $\Delta \eta$ is related to $M_x$ via the relation $\Delta \eta = -\ln \left( \frac{M_x^2}{s} \right)$. Since it is supposed that quantum numbers are not exchanged between the colliding protons, this leads to study diffractive processes in terms of Pomeron exchange. In hard diffractive processes the Pomeron is identified as a colorless gluon ladder exchanged by partons. However, there is not yet a satisfying theory which can explain all the aspects of this kind of hadronic processes and the understanding of the nature of
2.3 The diffractive process

![Diagram of diffractive processes]

Figure 2.5: Typical event topology in the $\eta - \phi$ plane for non diffractive (Minimum Bias) and diffractive processes (SD, DD, DPE, Multi Pomeron processes). The relative cross section values, measured at Tevatron and expected at the LHC respectively, are also reported on the right column.

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross Section (Minimum Bias)</th>
<th>Cross Section (LHC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic Scattering</td>
<td>18 mb</td>
<td>~30 mb</td>
</tr>
<tr>
<td>Single Diffraction (SD)</td>
<td>9.4 mb</td>
<td>~10 mb</td>
</tr>
<tr>
<td>Double Diffraction (DD)</td>
<td>6.3 mb</td>
<td>~7 mb (for $\Delta\eta &gt; 3$)</td>
</tr>
<tr>
<td>Double Pomeron Exchange (DPE)</td>
<td>~1 mb</td>
<td></td>
</tr>
<tr>
<td>Single + Double Diffraction (SDD)</td>
<td>1.5 mb</td>
<td>~3 mb (for $\Delta\eta &gt; 3$)</td>
</tr>
<tr>
<td>Multi Pomeron Exchange</td>
<td>~65 mb</td>
<td></td>
</tr>
</tbody>
</table>
Pomeron interactions is still a big challenge. Another important feature of diffractive events is that the majority of them show “leading” protons in the final states. These leading protons appear intact from the interaction region and are characterized by their $t$ and by their fractional momentum loss $\xi \equiv \frac{\Delta p}{p} \sim \frac{M_x}{\sigma}$. According to the beam optics, these leading protons can be revealed with high efficiency by the TOTEM RP detectors. Even if TOTEM is able to measure $\xi$, $t$, and mass-distributions in soft DPE and SD events, in order to have the possibility to perform detailed studies of the full structure of diffractive events (with the optimal reconstruction of one or more sizeable rapidity gaps) the integration of TOTEM with the CMS detectors would be welcome. For this purpose the TOTEM triggers are designed to be also incorporated into the general CMS trigger scheme and common data taking [19] are foreseen in a later stage.
Chapter 3

Tuning of Geant4 simulation

Nowadays in high energy physics an important role is played by simulation. In fact, due to the increasing complexity of the experiments, simulating the detectors response is a fundamental tool in order to understand the physical processes under investigation. The simulation is especially useful to determine detection biases, measurement errors, background contribution and to estimate quantities which can be difficult to determine with a direct measurement. For these reasons it is important to have an optimal tuning of the simulation. In this chapter, after a brief introduction to the Geant4 simulation package, the work done, in order to have a better tuning of the simulation of interest for any analysis involving the T2 detector, is reported.

3.1 A brief introduction to Geant4

TOTEM is using the same software framework of CMS (CMSSW). CMSSW provides in particular an interface for various software tools like ROOT, Geant4, several MC generators (Pythia, Phojet, and so on..), the Iguana visualisation software and many others. In particular, Geant4 (acronym of GEometry ANd Tracking) [13] is a platform for the simulation of the passage of particles through matter using Monte Carlo methods. It’s based on object oriented programming (C++ language) and it is a software toolkit developed at CERN. At the heart of Geant4 there is an abundant set of physics models which handle the interactions of particles with matter across a very wide energy range, allowing to reproduce the detector geometry and response,
as well as the decay processes of secondaries. The physical layout of an experiment (including detectors and passive materials) can be so schematized in the Geant4 geometry files and it is used by the simulation software to “evolve” a given physics event from the MC generation level (“particle level”) to the signals observed in the detectors (“digitization”).

In the CMSSW framework the Geant4 geometry is reproduced via the xml based Detector Description Language (DDL) [20]. The idea behind this detector description is that the information about the detector can be represented with a tree in which the nodes represent certain detector parts, and the edges represent a “part of” relationship. Every node in the tree can be uniquely mapped to a part or to a collection of parts in the “real” detector. The detector description in CMSSW consists of a hierarchy of geometry descriptions, based on direct acyclic multi graphs (acyclic multi graphs can be unfolded into a tree). Furthermore, this hierarchy contains parameters that give a detailed description about elements within this geometry description. To prevent a “part explosion”, i.e. the replication for several times of the descriptions of identical parts of the detector, the graph contains two layers of de-
3.2 Geant4 cuts tuning

Due to the big mismatch between data (collected with the T2 inclusive inelastic trigger in low luminosity runs, where the pile-up probability is negligible) and simulation (inelastic processes from Pythia MC [10]) observed in a first study of the T2 detector occupancy (as shown in figure 3.2), it was clear the importance to check the goodness of the simulation of the volumes in the forward region, which could affect the T2 telescope response. A comprehensive study has then been performed in order to check the proper Geant4 implementation and tuning of the simulation of the Beam Pipe, of the CMS components of interest and of the T2 detector. While dealing with the T2 geometry simulation optimization in the next section, in the present one we focus our attention on two fundamental aspects of the CMS simulation tuning: the cut parameters and the list of included volumes. To have an idea of the most important volumes to simulate properly for the T2 detector in figure 3.1 the CMS [21] experimental apparatus with its subdetectors is shown. The cut parameters define ranges in the different materials, so that particles with a range lower than the cut parameters value are not generated in the simulation (which anyway properly takes into account their energy deposition). The cuts used in the simulation are basically of two types: the so called “default cut” and the “cuts per region”. The default cut affects all the simulated volumes; the cuts per region affect only some “selected” volumes (listed in appropriate files). All the volumes, representing the CMS and

scriptions. A “part-type” containing the material and solid shape information, and a collection of parts position. A part of the detector is described by a “part-type” and its position relatively to another part-type (called parent part-type). Furthermore, a part can have a “parameterised” position. This is a description of a part that permits to place it several times in a parent part-type. In the xml files a number denotes a parameterisation of the parts. Attached to every part-type will be information about its material and geometry. Moreover certain parts (or part-types) have specific parameters related to their particular role (for example a silicon sensor can have a parameter “number of strips”).
3.2 Geant4 cuts tuning

Figure 3.2: Average number of pad clusters in each plane of the H0 T2 quarter (detector Id from 0 to 9). Comparison between data (black diamonds), simulation with improved default cut without cuts per region (blue triangles) and the old default cut (red circles). The low occupancy of plane 1 observed in data is due to an electric short on it, whose effect is not included in the simulation.

TOTEM detectors, that will be simulated by Geant4 are specified in a “global” list. The optimization of these cuts is clearly very useful in order to save CPU time. On the other hand, the CMS volumes (detector components) can be included or not in the simulation (acting on the global list), depending on their impact on the T2 detector response. While it is necessary to check if all of the volumes of interest are included in the simulation, the optimization of this list is also important in order to save CPU time (by removing the ones not affecting T2).

To perform this study the attention is focused on some interesting observables to check if a change in cuts or volumes would have some effect on the T2 detector response. We decided to consider the average pad cluster multiplicity per each plane, being the T2 detectors more efficient and less subject to noise in the more granular pad channel readout. Moreover this observable is simple (and so less biased, e.g. by tracking or alignment effects) and is directly related to detector occupancy, hence to the simulation response. Only the results related to the H0 (“plus near”: west, internal ring location) T2 quarter are report here; similar results have been obtained when considering the other three quarters. A proper tuning of the model reproducing the digital response of the detector, in terms of strip and pad cluster
efficiency and size distribution (“Digitization”), has already been performed using test-beam and Ion Collision data [22]. The goal is to find a “stable” configuration that uses the minimum CPU time, i.e. to reach a condition in which adding more volumes or pushing the cuts to a lower value does not have any substantial effect on the detector response. To achieve this result a lot of comparison between different simulation are shown in various plot. To ensure the best possible comparison all the simulation in a same plot are made with the same generation seeds, in order to avoid discrepancies due to difference in the physical events simulated. Anyway could happen that some events crash in a simulation and not in another. When is possible we correct on this crashed events, removing manually the corresponding ones from the simulation that don’t have the crashes, but is not always completely possible, so some little difference in the number of events between the two simulation could occur.

3.2.1 Geant parameters tuning

The simulation considered until now was using the default cut set to 1000 mm, defining the sampling range in the material, and the cuts per region set to values depending on the “selected” volumes of interest: all TOTEM detectors plus the beampipe from \( z = 0 \) m up to \( z = 6.5 \) m (set to 1 mm); the forward shielding around the T2 telescope (sampling set to 10000 mm). The results obtained with this simulation, which we will call ‘old default’ selection, are shown in figure 3.2, where they are compared with the data and with a simulation performed by setting the default cut to 0.01 mm without applying any cut per region. It seems clear that there is a huge difference between data (black diamonds) and simulation (red circles) that we can considerably reduce by moving the cut in the whole CMS volume to a lower value (blue triangles). However, as expected, this setting increases greatly the CPU time used in the simulation. So, for an optimal tuning also in terms of timing, it is important to understand what are the regions of importance for T2, so to reduce the cut value only there. A first check in this direction is reported in figure 3.3, where the red circles represent the results with the “old default cut” while the blue triangles represent the same situation in which we change only the value
3.2 Geant4 cuts tuning

Figure 3.3: Comparison between data (black diamonds), default sampling cut set to 1000 mm with cuts per region set to 0.01 (blue triangles) and the “Old cuts” (red circles). The low occupancy of plane 1 observed in data is due to an electric short on it, whose effect is not included in the simulation.

of the cuts per region parameter, setting it to 0.01 mm. We can see a little increase in the average pad cluster multiplicity, but the difference between simulation and data (black diamonds) is still huge. This means that there are some volumes not yet

Figure 3.4: Comparison between data (black diamonds) and default cut set to 1000 mm with cuts per region set to 0.01 mm, with two different volumes choices for the cuts region (red and blue markers). The cuts region represented by the red circles includes also the beampipe from 6.5 to 16 m, the beam radiation monitors and Castor. The low occupancy of plane 1 observed in data is due to an electric short on it, whose effect is not included in the simulation.
considered in the “selected” region which are really important for our detector. So we started to add additional volumes of potential interest, by including in the cuts region list the missing section of the beampipe (from 6.5 m up to 16 m), the beam radiation monitors and the Castor calorimeter. The result of this improvement is shown in figure 3.4 (red circles), where it is also compared with the data (black diamonds) and the previous situation in which no additional volume was included in the cuts per region list (blue triangles). We can see that this addition of volumes has a significant effect on the simulation and that now the result of using low cuts (0.01 mm) only for the volumes of interest is quite similar to the one obtained by applying low cuts to the whole CMS detector (setting the default cut to 0.01 mm). This is shown in figure 3.5, where the result obtained with only the default cut set to 0.01 mm (blue triangles) is directly compared with the one obtained with default cut set to 1000 mm and cuts per region set to 0.01 mm for the “improved” selected volume list (red circles).

Figure 3.5: Comparison between data (black diamonds), simulation with default cut set to 0.01 mm (blue triangles) and simulation with default cut set to 1000 mm and cuts per region set to 0.01 mm on the “improved” list (red circles). The low occupancy of plane 1 observed in data is due to a electric short on it, whose effect is not included in the simulation.
3.2.2 Getting the correct list of selected volumes

From figure 3.5 we can see that there is a residual mismatch between data and simulation, indicating there is still some important volume not included in the proper lists. After a first check, we found for instance that the HF calorimeter (CMS Hadron Forward calorimeter) is not included at all in the simulation, i.e. in the global volumes list that defines the simulated geometry of CMS. We then included HF, both in the global volumes list and in the cuts per region list and redo the simulation. The results are shown in figure 3.6, where the red circles represent the previous situation, the black diamonds are related to only the default cut set to 0.01 mm with HF included in the geometry and the blue triangles describe the inclusion of HF in the cuts per region list (and obviously in the geometry too). With HF included we have a decrease in the average pad cluster multiplicity, evidently as a consequence of some shielding effect on T2, and an increase in the difference between data and simulation.

![Figure 3.6: Comparison between the default cut set to 1000 mm with cuts per region set to 0.01 mm without HF (red circles) and with HF included (blue triangles). The black diamonds show a simulation with only default cut set to 0.01 mm and HF included.](image)

From this example it is clear that it is important to properly check if all the CMS volumes potentially of interest are included in the simulation. As first step the default cut is set to 0.01 mm (no cuts per region) for *ALL* the CMS volumes and the obtained result is used as “benchmark” to properly tune the volume list.
and parameter value for the cuts per region option. As usual, the reason for this additional tuning is the minimization of the CPU time required for the simulation. Following this perspective we then added to the global volume list other volumes still not included: CMS tracker, calorimeters and muon detectors. The results are shown in figure 3.7, where the black diamonds represent the “benchmark” situation and the blue triangles the one with the new global geometry, the default cut set to 1000 mm and the cuts per region set to 0.01 mm (with the selected volume list unchanged). The red circles instead represent the situation with the old geometry, the default cut to 0.01 mm and no cuts per region. The comparison between black diamonds and red circles markers shows that some volumes in the previous simulation were missing, while the comparison between black diamonds and blue triangles markers indicates that in the new simulation some volumes are missing in the cuts per region list.

We then included in the cuts per region list also the CMS central region (tracker, calorimeters, muon detectors). Because of software problems (a big amount of crashes) arising when applying weak cuts (like 0.01 mm) in the whole volume of these additional CMS detectors, we used for them some cut lists already defined in-

Figure 3.7: Simulation with only the default cut set to 0.01 mm with all the CMS volumes included (black diamonds) compared to a simulation on the same geometry and default cut set to 1000 mm, cuts per region set to 0.01 mm (red circles). And with a simulation (blue triangles) with default cut to 0.01 (no cut per region) and the old geometry (no CMS tracker, calorimeters and muon detectors).
side the CMS software framework. These lists have different cuts value for different part of the detectors, but on average the value is about 1 mm. The plot of figure 3.8 shows the comparison between the usual benchmark simulation represented in black diamonds and the new one depicted in blue triangles. From this plot it can argue that there is some volume in the central region of CMS that gives some contribution to the pad cluster multiplicity in T2. In fact now there is no more difference between the benchmark simulation and the new one with the use of cuts per region, meaning that now it is included some important volume that was excluded in the previous simulation. It was found that the origin of this contribution is the CMS Hadronic calorimeter “Hcal”, as shown in figure 3.9. Here the benchmark simulation (black diamonds) gives similar results with respect to another simulation obtained by adding only Hcal to the usual selected volume list (beampipe, forward shielding, Castor and TOTEM detectors) to which the cuts per region are applied. It is very important to notice that this choice of cuts, besides getting the same results, allows us to reduce the CPU time of about 30% with respect to the benchmark simulation in which we use only low default cuts on all CMS volumes. In order to further reduce the CPU time, we tried to remove other volumes (potentially not contributing to
3.2 Geant4 cuts tuning

Figure 3.9: Comparison between the simulation with only default cut set to 0.01 mm on all CMS volumes (black diamonds) and another (blue triangles) with default cut set to 1000 mm and cuts per region applied to the usual selected volumes (beampipe, HF, forward shielding, Castor and TOTEM detectors) and in addition to the CMS Hadronic calorimeter. The cuts per region value is 1 mm on average for the calorimeter and 0.01 mm for the other volumes.

the T2 signal) from the global volume list. The result is reported in figure 3.10, showing that there is no difference in including or not these volumes (Ecal, tracker and muon detectors) in the global CMS simulation. Unfortunately, we found that there are not appreciable differences in the corresponding CPU time usage. So we finally decided to always use for the simulation the “complete” geometry scheme, in order to be more conservative. The CPU time is much more affected by the choice of the cut parameter values, as it will be discussed in the next section.

3.2.3 Geant parameter optimization in terms of CPU time

After achieving a stable configuration for what concerns the cuts per region and the list of volumes of interest for T2, then the attention is focused on the cut parameter values. We wanted to understand if there is the need to decrease them (so that the simulation could better describe the data) or if there is the possibility to increase them (so to further reduce the CPU time needed for the simulation). This study was then done by reducing and by increasing the cut values. Figure 3.11 represents the comparison between the usual benchmark simulation (black diamonds) and a
3.2 Geant4 cuts tuning

Figure 3.10: Comparison between two simulations with the same cuts (default cut set to 1000 mm, cuts per region set to 0.01 plus proper cuts on Hcal) with different global volumes included in the geometry. The black diamonds have all CMS volumes included, the blue triangles only the volumes of interest (TOTEM detectors, Castor, beampipe, HF, forward shielding and Hcal).

Simulation with the cuts per region lowered to 0.005 mm (blue triangles). It is clear that there is not any further improvement, so it is not convenient to use looser cuts because it only increases the CPU time without any advantage. Figure 3.12 shows the comparison between the benchmark simulation (black diamonds) and a simulation with the cuts per region increased up to 0.1 mm. Here some decrease in
the average pad cluster multiplicity is shown, but on the other hand this cut choice provides a reduction of $\sim 45\%$ in the CPU time used with respect to the benchmark simulation. In conclusion, the best choice for a more realistic simulation, together with a convenient reduction of CPU time usage, is the configuration with the cuts per region set to 0.01 mm, giving a reduction of $\sim 30\%$ in the CPU time used with respect to the benchmark simulation.

Figure 3.12: Comparison between the default cut set to 0.01 mm (black diamonds) and the default cut set to 1000 mm plus cuts per region set to 0.1 mm (blue triangles).

Finally we have reached a stable setting for the simulation, in which the mismatch with data is not anymore related to a wrong use of the Geant4 settings and the CPU time usage is also optimized. Furthermore, all the interested volumes of the CMS detector for the production of secondaries in T2 are now understood. The missing ones are included in the simulation and the proper cuts for all of them are set. So, what is still missing in order to have the simulation completely under control is to check the implemented geometry for the forward region, that during this work for the Geant4 cuts tuning we found to be defective. In particular checking on: the beampipe, the ion pumps, the shielding before and around T2 and for the T2 telescope itself (cooling system and the support structure) are needed. Moreover it is important to understand the noise in the data (contributions and features) and, if necessary, to properly include it in the simulation. The optimization of the forward region geometry is discussed in the next section, while the noise is treated in the next
chapter of this thesis work. In the following chart, all figures and their contents are reported, in order to summarize the configurations of the various simulations done:

<table>
<thead>
<tr>
<th>Figure</th>
<th>Marker</th>
<th>Cut per region (mm)</th>
<th>Default cut (mm)</th>
<th>Cut per region list</th>
<th>Simulated Volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>blue triangle</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.2</td>
<td>red circle</td>
<td>1; 10000</td>
<td>1000</td>
<td>TOTEM, beampipe up to 6.5 m (1); Shield (10000)</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.3</td>
<td>blue triangle</td>
<td>0.01</td>
<td>1000</td>
<td>TOTEM, beampipe up to 6.5 m; Shield</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.3</td>
<td>red circle</td>
<td>1; 10000</td>
<td>1000</td>
<td>TOTEM, beampipe up to 6.5 m (1); Shield (10000)</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.4</td>
<td>blue triangle</td>
<td>0.01</td>
<td>1000</td>
<td>TOTEM, beampipe up to 6.5 m; Shield</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.4</td>
<td>red circle</td>
<td>0.01</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor.</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.5</td>
<td>blue triangle</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.5</td>
<td>red circle</td>
<td>0.01</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor.</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.6</td>
<td>blue triangle</td>
<td>0.01</td>
<td>0.01</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF</td>
<td>TOTEM, Beampipe, Shield, HF.</td>
</tr>
<tr>
<td>3.6</td>
<td>red circle</td>
<td>0.01</td>
<td>0.01</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor.</td>
<td>TOTEM, Beampipe, Shield.</td>
</tr>
<tr>
<td>3.6</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, Beampipe, Shield, HF.</td>
</tr>
<tr>
<td>3.7</td>
<td>blue triangle</td>
<td>0.01</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.7</td>
<td>red circle</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, Beampipe, Shield, HF.</td>
</tr>
<tr>
<td>3.7</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.8</td>
<td>blue triangle</td>
<td>0.01; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.01); CMS muon, tracker, calorimeters (1).</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.8</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.9</td>
<td>blue triangle</td>
<td>0.01; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.01); CMS Hcal (1).</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.9</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.10</td>
<td>blue triangle</td>
<td>0.01; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.01); CMS Hcal (1).</td>
<td>TOTEM, Beampipe, Shield, HF, Hcal.</td>
</tr>
<tr>
<td>3.10</td>
<td>black diamonds</td>
<td>0.01; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.01); CMS Hcal (1).</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.11</td>
<td>blue triangle</td>
<td>0.005; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.005); CMS Hcal (1).</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.11</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.12</td>
<td>blue triangle</td>
<td>0.1; 1</td>
<td>1000</td>
<td>TOTEM, beampipe up to 16 m; Shield, BRM, Castor, HF (0.1); CMS Hcal (1).</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
<tr>
<td>3.12</td>
<td>black diamonds</td>
<td>NO</td>
<td>0.01</td>
<td>NO</td>
<td>TOTEM, All CMS detectors.</td>
</tr>
</tbody>
</table>
3.3 Optimization of the simulated geometry in the forward region

As mentioned in the previous section, during the Geant4 cuts tuning, big mismatches between the simulated geometry and the real one in the forward region have been found. In particular, we found discrepancies in the description of: the T2 detector itself (supports and cooling system); the beampipe; the ion pumps before T2; the shielding around the detector. In this section we describe the work done in order to properly reproduce the geometry in the forward region, with the only exception of the beampipe which will be checked by experts of the CMS Collaboration. In order to have an idea of the most important volumes to simulate for the T2 telescope, we have to refer back to figures 3.1 and 1.2 where the TOTEM detector components inside CMS are depicted in a schematic view.

3.3.1 Geometry mismatches and simulation improvements

The geometry has been checked directly on the xml files, by comparing the dimensions and the positions of the various elements with CAD projects and also pictures. In this work an important aid is given by the “iguana” visualisation tool. A software that allows the visualisation of the detector geometry described by the xml files. This makes easier to compare real and simulated geometry and also to check the changes made on xml files. The main issues discovered with this work are: a cylinder of 1 mm steel around the T2 GEM, that isn’t there in reality; 4 ion pumps instead of 3 (and in a wrong position); bad dimensioned T2 horizontal supports; absence of T2 vertical supports (2 cm steel in front of Castor calorimeter); wrong dimension of the shielding around T2; a wrong gas mixture filling the T2 GEM and the absence of T2 VFAT FE electronics and HV cooling. All these mismatches have been corrected step by step, in order to understand the contribution of the single elements and hence its significance in the simulation. As for the previous cuts tuning, the simulations compared in the plots have been obtained with the same seeds and the same numbers of events. Moreover the observable depicted is still the
3.3 Optimization of the simulated geometry in the forward region

Figure 3.13: Simulated geometry for T2 and ion pumps before our corrections.

Figure 3.14: Comparison of pad cluster multiplicity between the simulation before any correction (blue triangles) and after the cylinder removal (black diamonds).

pad cluster multiplicity versus the detector plane number. The first step consists in removing the steel cylinder around the T2 GEM detectors. In figure 3.13 (a) is shown the T2 geometry before any modification, while in figure 3.21 (a) is depicted its geometry after the cylinder removal. The effect of this correction is visible on figure 3.14, which shows a comparison between the simulation before the correction (blue triangle) and the simulation after the cylinder removal (black diamonds). It is clear that this improvement has basically no effects. The next step has been a partial correction of the ion pumps geometry. In particular 1 pump has been removed, while the remaining 3 have been moved to the correct x-y position and tilted around z in the proper way. The simulation response before this ion pumps correc-
3.3 Optimization of the simulated geometry in the forward region

Figure 3.15: Comparison of pad cluster multiplicity between the simulation before (blue triangles) and after (black diamonds) a first correction on the ion pumps.

Comparing (blue triangles) and after (black diamonds) is shown in figure 3.15. From the comparison, it results that this change in the geometry has a not negligible effect on the simulation; in particular removing one pump and tilting the others leads to a little increase of the activity in the detector. In figure 3.13 (b) the pumps geometry before the changes is shown, while in figure 3.16 (a) it is displayed the geometry after these first corrections. Moreover in figure 3.17 (b) there is a picture of the flange with the ion pumps. Is important to notice that in the picture the ion pumps are in the “close” configuration, but it is possible to “open” them by 11 degrees, in order to move them partially outside the T2 geometrical acceptance. After this first correction on ion pumps geometry, another great improvement has been obtained in the geometry description of the shielding enveloping T2. The previous description of this shielding had a wrong size of the inner radius, in particular the radius is 33 cm instead of the previously quoted 25 cm. Therefore, it came out that the shielding is narrower than in reality and that in the simulation there was too much distance between the detector and the steel of the shielding. In figure 3.18, the comparison between the simulation before (blue triangles) and after (black diamonds) the shielding correction is shown. It is easy to notice that this modification leads to a reduction of the pad cluster multiplicity, as expected because of the increasing of the shielding material. The modifications made on the geometry are shown in
3.3 Optimization of the simulated geometry in the forward region

Figure 3.16: The ion pumps partially corrected (a) and in their final configuration (b). The beam-pipe tube, flange and T2 detector are also visible.

Figure 3.19, where pictures (a) and (c) show the old geometry of the shielding. The grey/white solids represent the shielding and the blue ones the T2 volume, in which are included the detector itself, the supports, the cooling system etc.. Figures (b) and (d) instead show the corrected geometry, with the inner radius of the shielding reduced and the T2 volume decreased. That means that now there is less space between the detector and the shielding. Another improvement has been the correction of the T2 horizontal (cylindrical) support dimensions and the implementation of the vertical supports. Concerning the horizontal supports, figure 3.20 shows the contribution of this modification in terms of pad cluster multiplicity (as usual). Looking at the plot, it seems evident that reducing the support dimensions, from the old 18 mm radius and 8 mm thickness to the actual 9 mm radius and 1 mm thickness, the result is a little increase of the multiplicity. In fact in figure 3.20 the blue triangles represent the simulation with the wrong supports and the black diamonds the one with the correct supports. This behaviour can be explained thinking to the greater shielding capability of the “old” bigger supports, with respect to the correct ones. This hypothesis is tested in figure 3.22 in which two simulations, with wrong
3.3 Optimization of the simulated geometry in the forward region

(a) T2 quarter, minus side.

(b) Beam-pipe flange and pumps.

Figure 3.17: One arm of the T2 detector (a), and the beam-pipe flange with the ion pumps (b). The T2 arm is the minus one, and the two quarters that compose it are partially opened. The ion pumps are shown in their closed position.

Figure 3.18: Comparison of pad cluster multiplicity between the simulation before (blue triangles) and after (black diamonds) the correction of the geometry of the shielding around T2.
Figure 3.19: Comparison between the old shielding geometry and the new one. Pictures (a) and (c) represent the “old” shielding (gray/white) and the T2 volume (blue) in two different views. Pictures (b) and (d) represent the “new” shielding (grey/white) and the T2 volume (blue).

Figure 3.20: Comparison of pad cluster multiplicity between the simulation before (blue triangles) and after (black diamonds) the correction of the T2 horizontal supports geometry.

(black diamonds) an correct (blue triangles) supports, are compared, but removing in both cases the main sources of secondaries from the geometry (HF and the shield-
ing around the detector). Since in this plot there are essentially no differences in the two cases (because there are not secondaries to shield), this supports the hypothesis that actually the wrong supports could be responsible of a shielding effect. The new implemented geometry for the supports is shown in figure 3.21. Picture (a) shows the bad dimensioned supports, while picture (b) displays the corrected horizontal supports and also the vertical supports implemented from scratch. For what concerns the addition of these vertical 2 cm steel supports, we have to notice that the unusual shape (rounded edges) is needed in the simulation scheme to avoid overlapping with the shielding. In the reality these supports are simple steel bars, also visible in the picture of the T2 minus arm in figure 3.17 (a). The effect of adding these steel “bars” is shown by figure 3.23 where the simulations with vertical supports (black diamonds) and no vertical supports (blue triangles) are compared. As expected, this modification had substantially no effect on our detector (these steel bars are in fact on the other side of T2 with respect to the IP), but could have some importance for the Castor calorimeter, which is located just behind T2.

A further improvement achieved is the final correction of the ion pumps simulated geometry. As mentioned before, a first correction on this detail has not been completely satisfying, because the pump dimensions, the z position and the angle (with respect to the beam-pipe) still did not correspond to the real geometry. So we applied the last correction on this part of the geometry. The final ion pumps geometry description is shown in figure 3.16 (b), and we can compare it to the pumps geometry after the first stage correction in picture (a). The effect of this change
Figure 3.22: Comparison of two simulations without HF and the shielding (enveloping T2) with the T2 horizontal supports geometry corrected (blue triangles) and none (black diamonds).

is shown in figure 3.24, the comparison between a simulation with the ion pumps at first stage correction (blue triangles) and another with the final corrected pumps (black diamonds) is made. This improvement seems to have quite no effect, but anyway the simulated geometry is now closer to the real one. The same thing can be said for the cooling system implementation and the gas mixture correction. For what concerns the cooling, this feature was not simulated at all. So we decided to schematize it with a horseshoe shaped 6.6 mm thick (18 mm large) aluminium piece (surrounding the detector planes) and a circular (8 mm thick, 24 mm in radius) aluminium piece. These two pieces are placed in the right configuration with respect to the T2 detector plane, as shown in figure 3.25 in which the plane with (b) and without the cooling (a) is shown. Moreover the T2 detector with the complete cooling system is shown in figure 3.26 (b), compared to the T2 detector without it (a). The result of the cooling implementation instead is shown in figure 3.27 where the simulation without the cooling system (blue triangle) is compared to the simulation with that feature (black diamonds).

The gas mixture correction is the last change made in the simulation. In fact the gas flowing through the T2 GEM planes was simulated as a mixture of Ar(80%)-CO\textsubscript{2}(20%) instead of the real mixture employed, which is Ar(70%)-CO\textsubscript{2}(30%).
3.3 Optimization of the simulated geometry in the forward region

Figure 3.23: Comparison of two simulations with (black diamonds) and without (blue triangles) T2 vertical supports.

Figure 3.24: Comparison of a simulation with a first stage ion pumps correction (blue triangles) and another with the final corrections (black diamonds).
3.3 Optimization of the simulated geometry in the forward region

Figure 3.25: (a) T2 detector plane without the cooling, (b) same plane with the aluminium cooling added.

Figure 3.26: The entire T2 detector complete of supports and cooling system is shown in (b), while in (a) the same detector is visible without the cooling.

effect of this further correction is shown in figure 3.28, where the simulation with the old gas mixture (blue triangles) and with the new correct one (black diamonds) are compared. After this last change we can consider the geometry optimization concluded. We can notice that some changes like gas mixture, cooling system, vertical supports have quite no effect on simulation. While other i.e. the horizontal supports, the ion pumps and especially the forward shielding have non negligible affects on the simulation. Anyway since also the non relevant changes has been implemented and yield (at least ideally) the geometry more similar to the real one we decide to keep on with these correction in the simulation.
3.3 Optimization of the simulated geometry in the forward region

Figure 3.27: Comparison of simulation with (black diamonds) and without (blue triangles) the T2 cooling system.

Figure 3.28: Comparison between the old gas mixture Ar(80%)-CO$_2$(20%) simulation (blue triangles) and the new simulation (black diamonds) with the correct mixture Ar(70%)-CO$_2$(30%).
3.4 Comparison between data and the improved simulation

After all the optimization made in the two previous section: Geant4 cuts parameter values, list of volumes of interest for the T2 response, geometry of the forward region. We want now to make a further comparison between data (T2 inclusive inelastic trigger on low luminosity runs) and simulation (inelastic processes from Pythia). In order to check the best achievable result with the current implementation of the simulation. For this reason this study has been made by using an updated implementation of the clusterization, tracking and digitization algorithms. Figure 3.29 shows the comparison between data and the best currently available simulation, for the average pad cluster multiplicity for each plane of the H0 T2 quarter. In particular we can notice that the low occupancy of plane 1 observed in data, due to an electric short on it, is now well reproduced by the simulation (as compared for instance with Figure 3.2), but there is still a not negligible global mismatch between data and simulation characterized by a $\sim 30\%$ higher activity in the data.

![Figure 3.29: Comparison between data and simulation with the best tune currently available (i.e. best choice of Geant4 cuts and of CMS volumes, checked geometry) and with the new digitization, cluster and tracking algorithms.](image)

We decided to further investigate the origin of this residual mismatch. Figure 3.30 (top) shows the comparison between data and simulation for the average number
of pad clusters in plane 0 of the H0 T2 quarter, as a function of the number of reconstructed tracks in the whole T2 detector. Now, by relating the response of a T2 quarter with the total “activity” in the detector, the simulation can substantially reproduce reasonably well the data (which are anyway characterized by a higher average track multiplicity, see figure 3.33). As a control check, as shown in figure 3.30 (bottom), a direct comparison between Pythia and Phojet [12] MC generators has been performed showing a very good agreement. Figure 3.31 shows the number

![Figure 3.30](image)

Figure 3.30: Average pad cluster multiplicity in plane 0 of the H0 T2 quarter versus the number of reconstructed tracks per event in the whole T2 detector. Top: comparison between data and simulation (Pythia). Bottom: comparison between simulations made with Pythia and Phojet MC generators. The most updated simulation and the new digitization, clusterization and tracking algorithms have been used.
of pad clusters for plane 0 (left) and 9 (right) of quarter H0 for both data (top) and simulation (bottom). These plots assure that the mismatch between data and simulation are not due to strange tails in the pad cluster number distributions. Figure 3.32 allows to compare the pad cluster size on planes 0 and 9 of the H0 quarter. For this parameter there is a good agreement between data and simulation, indicating that in both cases these clusters are really due to particles hitting T2 and are not biased by other effects, like detector noise. This assumption has to be properly tested with appropriate noise studies performed on data, from which the typical noise signal properties have to be understood. Figure 3.33 shows the number

![Figure 3.31: Number of pad clusters for plane 0 (left) and 9 (right) of quarter H0 for data (top) and for simulation (bottom). The most updated simulation and the new digitization, clusterization and tracking algorithms have been used.](image)

of all tracks and of primary tracks reconstructed in the T2 detectors, both for data (top) and simulation (bottom). A track is defined “primary” when passing standard cuts \( Z_0 < 6000 \text{ mm}, \ R_0 < 40 \text{ mm} \) and \( \chi^2_{prob} > 0.01 \), for more detail on these parameters see section 1.3) on the compatibility of the track origin with the event
Figure 3.32: Pad cluster size for plane 0 (left) and 9 (right) of quarter H0 for data (top) and for simulation (bottom). The most updated simulation and the new digitization, clusterization and tracking algorithms have been used.

vertex. The observed discrepancy between data and simulation can be attributed to the fact that the MC generator is not a priori reproducing well the inelastic processes, as well as to the fact that there is still some part of the CMS apparatus not well simulated, which produces secondaries hitting T2. These hypotheses are supported by the observation that the simulation predicts, comparing to data, a lower average global track multiplicity (mainly due to secondary particles produced in the material in front of or around T2), against a higher average primary track multiplicity (mainly due to particles generated at the interaction point). In order to have an idea of how much of this mismatch can be attributed to the physics itself, we compared the results obtained with Pythia and Phojet. As shown in figure 3.34, a ∼ 10% difference in average pad cluster multiplicity was found. That is not enough to explain the difference between data and simulation.

### 3.4.1 Conclusions and possible explanation

The previous comparison between data and simulation points out that the tuning of the simulation leads to appreciable results, in optimizing the simulation with
3.4 Comparison between data and the improved simulation

Figure 3.33: Number of all tracks (left) and primary tracks (right) reconstructed per event in the whole T2 detector in data (top) and in simulation (bottom). The most updated simulation and the new digitization, clusterization and tracking algorithms have been used.

Figure 3.34: Comparison of the average pad cluster multiplicity in the H0 T2 quarter, as obtained with Pythia and Phojet MC generators. The most updated simulation and the new digitization, clusterization and tracking algorithms have been used.

respect to the data. Since looking back to the red circles of figure 3.2, representing the starting point of our work, and comparing them with the blue triangles of 3.29, we can see a good enhancement of simulation to approach the data. Anyway there
is still an absolutely non negligible residual difference between observed data and simulation. Especially looking at the average pad cluster multiplicity per plane, this is on average about 24% less in simulation with respect to the data. This discrepancy could not be attributed to the uncertainty in the generation model; in fact the simulation made with two different generators, Pythia6 and Phojet (figure 3.34) points out an uncertainty of about 10% on the average pad cluster multiplicity per plane. Anyway this is for sure a relevant issue for the simulation and an improved tune of the generator could produce appreciable effects and has to be taken into account. Moreover one of the possible causes at the origin of this discrepancy had to be searched in possible problems related to the propagation of the simulated particle through the different volumes. In fact it is known that due to a problem in the CMSSW version used at the moment by the TOTEM offline SW, related to the shower modeling in HF, the particles that reach the inner side of HF are no further propagated by Geant4. This problem will be solved by the migration of the TOTEM offline software to a newer version of CMSSW (migration that is ongoing), but at the moment is present and no other way to solve it were known. To evaluate the importance of this issue we produced a Particle Gun, that on the contrary of the full simulation we know as not affected by this problem, by firing a particle directly on the problematic region $5.18 < \eta < 5.32$. We decided to fire one by one in this range both $\pi^+$ and $\gamma$, and in figure 3.35 four plots for this simulation are shown. The plots on the top show the mean number of pad clusters produced in the plane 0 of quarter H0 as a function of the $\eta$ of the fired particle. While the plots on the bottom show the same quantity but for the plane 9 of quarter H0. The plots on the left refer to $\pi^+$ and the ones on the right to $\gamma$. The energy of the fired particles is between 10 and 60 GeV. From these plots we can see that the region investigated is really a critical one for our detector. In fact a particle going in this zone produces on average about $2\div2.4$ pad cluster on the planes of the detector if it is a $\pi^+$ and about $5.2\div6.2$ if it is a $\gamma$. Together with the fact that for each event there is a mean of $0.5\gamma$ and $0.6$ charged particles going towards this problematic region (these ratios are calculated at generator level with the simulation), it definitely becomes critical for a good simulation of the T2 detector response. So, it seems likely that this could
3.4 Comparison between data and the improved simulation

![Four plots showing average pad cluster multiplicity in a plane of T2 quarter H0, with respect to the η of the Particle Gun fired particle. The plots on the top refer to plane 0 and the ones on the bottom to plane 9. Plots on the the left are made with π⁺ and plots on the right with γ.](image)

Figure 3.35: Average pad cluster multiplicity in a plane of T2 quarter H0, with respect to the η of the Particle Gun fired particle. The plots on the top refer to plane 0 and the ones on the bottom to plane 9. Plots on the the left are made with π⁺ and plots on the right with γ.

explain at least in part the remaining discrepancy between data and simulation. In principle another possible source of this difference between data and simulation could be provided by the noise, whose contribution in T2 doesn’t have a suitable simulation. This motivates the study proposed in the next chapter, in which the characterization of the noise contribution for this detector has been treated.
Chapter 4

Noise contribution in T2 detector

An important feature to be evaluated for the T2 detector is the noise contribution in data. This estimate would also allow a better understanding of the simulation correctness. Noise is actually not simulated a priori (with the only exception of the capacitive pad and strip noise). Therefore it is important to determine the necessity of a simulation improvement.

The first problem is that the output of the detector is digitalised and there is no way to distinguish between a hit generated by a charged particle and a hit due to electronic noise. The strategy adopted here consists in tagging as noise hits the not-associated-to-a-track hits. At least in principle, this feature can permit to distinguish noise hits from the others. Obviously from this unique feature it’s very difficult to determine the noise contribution, because a hit not associated to a track could be generated from a secondary particle for which a track is not reconstructed. However, this aspect could be checked by comparing data and simulation. In fact the simulation has (substantially) no noise contribution and could be useful to see the ratio of hits from secondaries that are recognised as noise (hereafter “noise hit” means always hit not associated to a track). Another strategy to reduce the incidence of “fake” noise hits is to analyse only events with a reduced incidence of secondaries. In particular we can apply some selection in the topology of the events in order to choose only “clean” ones for the study.

For this study three different types of data taken with the T2 detector have been used. In particular we use proton data taken either with a T2 inclusive inelastic
trigger or with a bunch zero trigger. Bunch zero trigger means that the event has been triggered on the timing of the bunch crossing at IP5, while T2 inclusive inelastic means that an event has been triggered with an AND of the bunch crossing timing and a trigger signal auto-generated by T2 (using a special coincidence chip). These data have been collected in October (2010) during some dedicated low luminosity protons run at $\sqrt{s} = 7$ TeV and $\beta^* = 3.5$ m, with 5 bunches and $\sim 3 \times 10^{11}$ protons per beam. The other kind of data used are instead ion data, taken in December (2010) during a heavy ions ($^{208}$Pb) dedicated run at a centre-of-mass energy per nucleons pair $\sqrt{s_{NN}} = 3.5$ TeV with a T2 inclusive inelastic trigger. These ions data, as known from [23], are characterized by having an high rate and a lot of events with a low track multiplicity. This is ascribable to electromagnetically induced processes, that have very large cross sections at LHC energies but generate very low multiplicities. For this reason, in these particular data, most of the events in our detector are very “clean”, with a few primary tracks and a very reduced secondary incidence. This means that they represent a “golden sample” which can be used for T2 noise (and other detector performance) studies.

All the plots presented in this chapter refer to the T2 quarter H0 only (that is the quarter in the plus near location), here used as our reference quarter. However the results have been later checked for the whole T2 detector and the conclusions hold true also for the other quarters (even if the plots concerning these quarters for simplicity and concision are not shown here).

4.1 The noise characterization

In figure 4.1 the comparison between ion data, proton data and simulation (of protons) is shown in a plot where the average number of noise class 1 hits is shown with respect to the plane number of quarter H0. This plot is made with 10000 events of simulated protons, 10000 events of proton data and 10000 events of ion data. The same plot, but this time for class 2 noise hits, is displayed in figure 4.2. In order to better understand the plots, it’s important to remember: that the number of the plane (also called plane Id) increases with the distance from IP5; that class 1 hits
4.1 The noise characterization

Figure 4.1: Comparison between ion data (red circle), proton data (black diamonds) and simulated protons (blue triangles) for events with only one track in T2 quarter H0. The average number of class 1 noise hits is shown for each plane Id, for the quarter H0.

means that are hits composed by both a strip and a pad cluster; that class 2 hits means that it is constituted by only one cluster (either strip or pad); that the most distant planes from IP5 will be often simply called “last” planes.

Concerning class 1 hits, the behaviour of proton data and simulation looks quite similar, while the mean number of noise hits (between proton data and simulation) is different as well as the behaviour and the mean of ion data. In fact ion data do not show the big rise of noise hits with the plane number, present in proton data and simulation. Looking at class 2 hits in figure 4.2 it is still possible to observe a growth in the noise incidence on the last planes, except for ion data. But in this case the response for simulation and protons is quite similar, though it is believable that real noise hits could be more likely second class hits. So the bigger mismatch between proton data and simulation for the first class noise hits with respect to the second class ones, seems to suggest that probably this mismatch rises up from a different incidence of secondary particles, more than from noise itself. Moreover the origin of “noise” hits from secondary particles could explain also the growth of the noise signal in the last planes. There actually is no reason for which planes more distant from IP5 would be more noisy. While, as a track is required to pass the threshold of 4 hits to be recognised by the detector, it’s plausible that incoming particles that
4.1 The noise characterization

Figure 4.2: Comparison between ion data (red circle), proton data (black diamonds) and simulation (blue triangles) for events with only one track in T2 quarter H0. The average number of class 2 noise hits is shown for each plane $I_d$, for the quarter H0.

cross only the last planes of the detectors (because they originate away from IP) are not tracked. For this reason, in order to try to reduce the contribution of secondary particles, the plots in figure 4.3 and 4.4 are made only for events with a track in T2 quarter H0 and no tracks elsewhere. The two plots are still obtained for quarter H0 and show the number of class 1 and class 2 (respectively) noise hits versus the plane $I_d$. Now the sample of simulation and proton data used for the study is increased to 100000 events (respect to the 10000 events used in the previous plots) to avoid problems of low statistics. In fact the new selection rule for the events reduces by an order of magnitude the number of selected ones in proton data and simulation, while having less severe effects on the ion data, and then the sample for ions has not been increased. It is important to notice that the change in the selection rules for the analysed events leads to a reduction of noise hits contribution by about the 50%, for simulation and protons, for both first and second class hits. On the other hand in ion data the noise contribution remains substantially the same. Moreover also the behaviour of all the data remains similar. This is in agreement with the hypothesis that in the plots of figure 4.1 and 4.2 the main contribution to “noise” came in reality from secondary particles. In fact by requiring more “clean” events the noise incidence decreases, both for data and simulation (it is important to remember
4.1 The noise characterization

Figure 4.3: Comparison between ion data (red circles), proton data (black diamonds) and simulation (blue triangles) for events with only one track in T2 quarter H0 and no tracks in the other quarters. The plot shows the average number of class 1 noise hits versus plane Id, for the quarter H0.

Figure 4.4: Comparison between ion data (red circles), proton data (black diamonds) and simulation (blue triangles) for events with only one track in T2 quarter H0 and no tracks in the other quarters. The plot shows the average number of class 2 noise hits versus plane Id, for the quarter H0.
4.1 The noise characterization

Figure 4.5: Comparison between ion data (red circles), proton data (black diamonds) and simulation (blue triangles) for events with no tracks in T2 arm plus (quarters H0 and H1) and only one track in T2 arm minus (quarters H2 and H3). The plot shows the average number of class 1 noise hits versus plane Id, for the quarter H0.

that the simulation has no noise implemented). This is not explicable in terms of electronic noise. Moreover the situation for ion data remains quite the same, because these data are cleaner since the beginning and the selection of the events is not so critical for them. From this perspective, in order to have a better understanding of the noise not induced by particles, we decided to make the same plots (of noise hits versus plane) with a even tighter selection of events. This selection, that allows to have events with less secondary particles, is achieved by requiring only one track in T2 arm minus and no tracks on T2 arm plus. The plots refers to quarter H0, as usually, which belongs to the arm plus. Therefore only events with only one track in the opposite side with respect to the quarter under investigation are analysed in figure 4.5 and 4.6. Figure 4.5 shows the number of class 1 noise hits, while figure 4.6 shows the class 2 noise hits. For both noise classes the behaviour is similar to the previous plots (figure 4.3, 4.4) but the values of the plot mean are reduced by about an order of magnitude. This points out that previous results were still depending on secondary particles crossing the detector. Certainly this effect is less significant in the last plots, even if probably it is not completely removed. Anyway, from these plots the incidence of noise seems really small. In fact, looking at the mean value of noise hits, we conclude that there is less than one noise hit (either class 1 or class
4.1 The noise characterization

Figure 4.6: Comparison between ion data (red circles), proton data (black diamonds) and simulation (blue triangles) for events with no tracks in T2 arm plus (quarters H0 and H1) and only one track in T2 arm minus (quarters H2 and H3). The plot shows the average number of class 2 noise hits versus plane Id, for the quarter H0.

2) per event in the whole quarter. This is sufficient to affirm that the detector is silent, at least when there are no charged particles crossing it. Obviously there is the possibility to have a noise contribution induced by the passage of charges, but this is really difficult to study, at least with these methods. Anyway, something on this direction has been done and it will be reported later on. Already at this point of the study the noise contribution for the T2 detector seems not that significant to be simulated, and it could not be the source of the discrepancy between simulation and data found in chapter 3. Anyway we decided to have another look to the noise by using a different approach. In particular, the noise contribution is now extrapolated analysing a run of proton data with bunch zero trigger (that means that the detector is triggered on the timing of the bunch crossing), looking at hits in events with no tracks. This way allows us (at least in principle) to look at the detector in a data acquisition configuration, but when there are no particles crossing it. So that is probably the best way to study the “pure” electronic noise of the detector (i.e. not induced by charged particles). Figures 4.7 and 4.8 show the results of this study, for first and second class noise hits respectively. In these plots the noise contribution is reduced by about two orders of magnitude (with respect to figures 4.5 and 4.6), again demonstrating that the “pure” electronic noise is not an issue for the T2 detector.
Figure 4.7: Average number of class 1 noise hits versus plane Id, for the quarter H0, obtained on proton data taken with a bunch 0 trigger. For this plot only events without any tracks are required.

Figure 4.8: Average number of class 2 noise hits versus plane Id, for the quarter H0, obtained on proton data taken with a bunch 0 trigger. For this plot only events without any tracks are required.
4.1 The noise characterization

Figure 4.9: Average number of noise hits versus number of tracks in the T2 quarter H0. Events with tracks in H0 only have been selected.

Moreover in figure 4.9 and 4.10 two plots to check the dependency of noise on the detectors activity are shown. The plot in figure 4.9 displays the number of noise hits with respect to the number of tracks in quarter H0. Also in this case some selection cuts were applied in order to have a minor contribution from secondary particles. In particular for this plot, only events with tracks in the H0 quarter are required. The plot is made as usual for proton data (black diamonds), proton simulation (blue triangles) and ion data (red circles). A strong correlation between number of noise hits and number of tracks is appearing. This could be interpreted as an increase of real electronic noise with the number of charged particles in the detector, but the simulation is following well the data, even without any noise implemented. For this reason it seems that the noise hits, also in this case, came mainly from secondary particles and not from the electronics. Therefore, even for a situation with many charged particles crossing the detector, the noise contribution doesn’t seem relevant. Notice that in figures 4.9 and also 4.10 the numbers of hits is summed over the 10 planes of the detector, while the previous plots were showing hits per plane. The dependency of noise hits from the number of tracks arises because in events with more tracks (i.e. more activity) there is also a greater incidence of non tracked particles. Instead the plot of figure 4.10 shows the number of noise hits in T2 quarter H0 versus the number of tracks in T2 arm minus (opposite to H0). In this plot there isn’t any
4.1 The noise characterization

![NoiseHitH0 MultiTrkVsNumTrkH2H3](image)

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Entries</th>
<th>Mean</th>
<th>Mean y</th>
<th>RMS</th>
<th>RMS y</th>
<th>Overflow</th>
</tr>
</thead>
<tbody>
<tr>
<td>DatiProt10^5</td>
<td>12609</td>
<td>3.641</td>
<td>0.6572</td>
<td>6.66</td>
<td>5.439</td>
<td>0</td>
</tr>
<tr>
<td>Simu10^5</td>
<td>12513</td>
<td>3.992</td>
<td>0.8151</td>
<td>7.021</td>
<td>4.838</td>
<td>0</td>
</tr>
<tr>
<td>DatiIoni10^4</td>
<td>12513</td>
<td>3.992</td>
<td>0.8151</td>
<td>7.021</td>
<td>4.838</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 4.10: Average number of noise hits in T2 quarter H0 relative to the number of tracks in the detector. Events with tracks only in T2 arm minus (the arm opposite to the one containing H0) have been selected.

The correlation between the number of tracks and the number of noise hits, but the mean value of the noise remains quite constant and the simulation describes the data quite well, without the need of a true noise simulation. In any case, some other plots have been made in order to have a better characterisation of the noise (or supposed one) and to see if there is some other feature that permits a better discrimination between noise hits and no noise ones. First of all, one interesting thing to be checked is the number of noise hits per plane in each event, because the previous plots show only the mean number of noise hits per plane and not how they are distributed. Figures 4.11 and 4.12 show the histograms of the number of first and second class noise hits for a single plane. Both the figures show a comparison between proton data (right side) and simulation (left side) for plane 0 of quarter H0 (top) and plane 9 of the same quarter (bottom). All the plots are made with the “tighter” selection on the events, that requires only a track in the opposite side with respect to the quarter analysed. Another important feature to be checked is the cluster size of strips and pads for the different hit types. The plots in figure 4.13 show a comparison of strip and pad cluster size for class 1 noise hits and “no noise” ones. The hits associated to a track are considered as no noise hits, because are more likely to be generated from the crossing of a charged particle than from electronic noise. The plots on the top
4.1 The noise characterization

Figure 4.11: Number of class 1 noise hits in plane 0 and plane 9 of quarter H0, for events with no track in T2 arm plus (quarters H0 and H1) and only one track in T2 arm minus (quarters H2 and H3). Histograms on the top refers to plane 0, the bottom ones to plane 9. Plots on the left are about the simulation and the ones on the right are about proton data.

refers to no noise hits, while the plots on the bottom refers to noise ones. Plots on the left are about strip cluster size and the ones on the right are about pad cluster size. Each plot itself shows a comparison between proton data (black line), ion data (red line) and proton simulation (blue line). For the no noise hits the pad and strip cluster size are quite similar for ions, protons and simulation. Especially ions are really close to the simulation, but this is not so strange because the simulation is tuned also on ion runs. Focussing on noise hits instead the strip and pad cluster size between each type is less similar and especially for protons and simulation is larger than for no noise hits. This could be caused by the larger contribution from hits generated by secondary particles, which have a skewer trajectory and so generate a hit bigger in size. Finally in figure 4.14 we analysed the data with bunch zero trigger, previously used to show the noise contribution. The strip and pad cluster size for the noise hits is shown under different conditions. On the top there are the plots concerning the first class noise hits, on the bottom the ones concerning the
4.1 The noise characterization

Figure 4.12: Number of class 2 noise hits in plane 0 and plane 9 of quarter H0, for events with no track in T2 arm plus (quarters H0 and H1) and only one track in T2 arm minus (quarters H2 and H3). Histograms on the top refers to plane 0, the ones on the bottom to plane 9. Plots on the left are about the simulation and the ones on the right are about proton data.

second class noise hits, while on the right there are the plots about the pad cluster size and on the left the ones about the strip cluster size.

After this study, it is now possible to conclude that the noise is not a relevant issue for the T2 detector, and there is no need to simulate it. In particular the largest fraction of hits not associated to a track can be attributed to secondary particles, which for some reason are not tracked. Especially from bunch zero crossings data (and ion collisions too), we can argue that the detector is really silent, at least for this situation of very low occupancy. For what concerns the noise induced by the passage of charged particles, this is more difficult to study and evaluate, but it seems anyway under control, as we can deduce by looking at figure 4.9.
Figure 4.13: Strip (top-left) and pad (top-right) cluster size for “no noise” class 1 hits (i.e. hits associated to a track) of quarter H0. Events with only one track in T2 quarter H0 (and whatever in the other quarters) are required. Strip (bottom-left) and pad (bottom-right) cluster size for noise class 1 hits. Events with no tracks in T2 arm plus (quarters H0, H1) and one track in T2 arm minus (H2 or H3) are required.
4.1 The noise characterization

Figure 4.14: Strip (left) and pad (right) cluster size for class 1 (top) and class 2 (bottom) noise hits for quarter H0, obtained on proton data taken with a bunch 0 trigger. Only events without any tracks are required.
Chapter 5

A first study of inelastic events
with the T2 detector

After a long and detailed study to optimize the detector simulation and to understand features like the detector noise and the contribution from secondaries, we finally moved to simulate inelastic proton-proton collisions and study the T2 response. We would like to verify in more detail the acceptance of the detector, previously studied at particle level only for the original TOTEM Technical Design Review, more than six years ago. As mentioned in the second chapter, a precise measurement of the total proton-proton cross-section requires a direct measurement of the inelastic rate, consisting in diffractive and non diffractive minimum-bias events, with the least possible loss of events by the acceptance coverage of the TOTEM inelastic detectors. The study presented in this section is important for the TOTEM Collaboration in view of the future dedicated data taking at low luminosity that, at the time of writing, unfortunately did not occur yet.

The measurement of the total hadronic cross-sections and their theoretical understanding have always been topics of crucial interest in particle physics. As they cannot be calculated by Quantum Chromo Dynamics (QCD), many phenomenological approaches have been used to describe the existing measurements. General arguments based on unitarity, analyticity, and factorisation imply a bound (the Froissart bound) on the high-energy behaviour of total hadronic cross-sections. This bound is independent of the details of the strong interaction dynamics and states that the
total cross-section can not rise faster than $ln^2(s)$, where $\sqrt{s}$ is the centre-of-mass energy. Recently it has been extended to the inelastic cross-section [24]. Existing experimental data show a rise in the hadronic cross-sections with $s$, but it is unclear whether the asymptotic behaviour has already been reached. The ATLAS experiment has recently published a paper on their first measurement of inelastic collisions at $\sqrt{s} = 7$ TeV [25], resulting in a cross-section of $60.3 \pm 2.1$ mb for $\xi > 5 \times 10^{-6}$, corresponding for diffractive events to require at least one of the dissociation masses to be larger than $15.7$ GeV. $\xi$ is the fractional momentum loss of the proton, $\xi = \Delta p/p$.

In the present section, we are going to show that in TOTEM, at the same center of mass energy, a limit almost one order of magnitude smaller than ATLAS could be reached thanks to the small angle coverage of the T2 detector.

Monte Carlo (MC) simulations are usually used to determine the acceptance of the event selection and to assess systematic uncertainties. The detector response to the generated events has been simulated using the TOTEM simulation based on Geant4. As for the ATLAS paper, the Pythia6 [10], Pythia8 [11] generators have been used to predict properties of the inelastic collisions.

Unfortunately, due to a bug found in the software interface between the Phojet [12] generator and the TOTEM offline analysis software, at the time of writing there was no chance to use Phojet for this kind of study. Because the momentum of the particles is not saved properly. The bug was reported to the TOTEM collaboration and will be corrected as soon as possible by the people expert on the subject.

These generators distinguish between different processes that contribute to inelastic pp interactions: single dissociative (SD) processes, $pp \rightarrow pX$, in which one proton dissociates; double dissociative (DD) processes, $pp \rightarrow XY$, in which both protons dissociate with no net color flow between the systems $X$ and $Y$; and nondiffractive (ND) processes in which color flow is present between the two initial-state protons. The model used by Pythia6 and Pythia8 predicts cross-sections of 48.5 mb, 13.7 mb and 9.3 mb for the ND, SD and DD processes, respectively. The cross-sections used by Pythia6 and Pythia8 are identical, but they differ in the modelling of the hadronic final state. The MC generators define the inelastic cross-section as the sum of these contributions, and thus Pythia (Phojet) predicts an inelastic
The variable $\xi$ is defined at the particle level by dividing the final state particles into two systems, $X$ and $Y$. The mean $\eta$ of the two particles separated by the largest pseudorapidity gap in the event is used to assign all particles with greater pseudorapidity to one system and all particles with smaller pseudorapidity to the other. The mass, $M_{X,Y}$, of each system is calculated and the higher mass system is defined as $X$, while the lower mass system is defined as $Y$. The variable $\xi$ is then given by $\xi = M_X^2/s$ and it is bounded by the elastic limit of $\xi > m_p^2/s$. Due to our limited detector acceptance, we want to study the $\xi$-range in which this measurement is restricted.

5.1 Study of Single Diffractive events

In this section the study of the detection efficiency and acceptance of the TOTEM inelastic telescopes T1 and T2 is presented for the SD events. For this purpose various samples of pure SD processes have been simulated from the generation (via Pythia6 and Pythia8 tools) up to the reconstruction level using the TOTEM dedicated software. The dependence of the efficiency on $\xi$ is pointed out, and the choice of a convenient $\xi$ cut value ($\xi_{vis}$), in order to define a kinematic range for the “visible” inelastic cross-section measurement, is explained. Moreover the factor $\varepsilon_{tot}$ that, relying on the simulation, allows to correct from the measured inelastic rate to the “real” one is calculated. For this kind of study the first important thing to do is to choose the criteria to identify an inelastic process (or event) as detected or not by the TOTEM telescopes T1 and T2. A logical choice is to define that an inelastic event is occurred when one of the two telescopes can reconstruct at least one track. In this way we are quite sure that there are no noise contribution in the detection of the events. However it could be that the requirement of a track is too much constraining, leading to loose a part of the inelastic processes. So we have to understand if this kind of request is the optimal one. Or instead if there are other possible better choices, that permit again to avoid noise contribution, but at the same time allow to detect more events. For this reason in figure 5.1 the $\xi$ distribution for various
5.1 Study of Single Diffractive events

Figure 5.1: $\xi$ distribution for: (a) all MC events; (b) events with at least one stable particle in T1-T2 range; (c) type 1 events; (d) type 2 events. All plots have been made with Pythia6 (black line) and Pythia8 (blue starred line), curves have been normalized to 1.

different selection criteria is reported. In (a) the distribution for all Monte-Carlo SD generated processes is shown. In (b) there are events that have at least one stable particle going toward the T1 or T2 telescopes (at MC level). In (c) instead there is the $\xi$ distribution for all processes that have at least one hit reconstructed in one of the two detectors (type 1 events). Finally in (d) is reported the histogram showing the $\xi$ of events with at least one track at reconstruction level in the detectors (type 2). Focussing simply on the number of entries of the four histograms, it’s easy to notice that the TOTEM inelastic telescopes are able to reconstruct about the 84.0-83.4% of the inelastic SD events (Pythia6-Pythia8 respectively), also requiring a track. This is only ~1.6-1.5 percentage points less than requiring one hit, that is clearly the minimal request we can make, but it can obviously be affected by noise.
Moreover referring to plot (b) we can argue that only in the $\sim 1.4\%$ of the cases a particle outgoing the T1-T2 range is not detected (with the criterion of one track). Finally it is not surprising that the number of events of type 1 (that pass the criterion of one hit) is higher than the ones with a particle in the detectors range. This in fact is well explainable in terms of secondaries, that could make revealable also processes without primary particle in the detector range. Since the request of one track causes a loss of events of about a percent, but, at the same time, minimizes the risk of noise, we decided that this “selection rule” is reasonably good for our purpose. Therefore we don’t investigate any further in search for a different criterion.

We now move on to investigate the detection efficiency for SD processes. This efficiency is clearly dependent on $\xi$, and to calculate it we decided to use the Bayesian approach [26, 27]. In fact it is well known that the two commonly used solutions for the errors calculation, Poissonian and Binomial approach, are both incorrect, because they lead to absurd results when brought to the limit cases. In particular, assuming that $n$ is the size of the sample, $k$ is the number of success (events that pass the selection criteria) and $\varepsilon$ is the “true” efficiency, the Binomial approach leads to zero errors in the limit cases of $k = 0$ and $k = n$. While the Poissonian errors calculation leads till to zero error in the case of $k = 0$ and to an unphysical error interval extent over 1 for the case $n = k$.

$$P(k; \varepsilon, n) = \binom{n}{k} \varepsilon^k (1 - \varepsilon)^{n-k}$$

(5.1)

For a correct treatment of the errors we can start again from the Binomial distribution (Eq. 5.1) thinking of $P(k; \varepsilon, n)$ as the probability that $k$ events will pass the cut. Given the condition that the efficiency is $\varepsilon$, that there are $n$ events in the sample and that our prior information is that the process is binomial. As we want to compute the errors on $\varepsilon$, what we need to determine is $P(\varepsilon; k, n)$. To calculate this
5.1 Study of Single Diffractive events

probability density function we use the Bayes theorem\(^1\) with the following ansatz:

\[ P(\varepsilon; k, n) = \frac{P(k; \varepsilon, n)P(\varepsilon; n)}{C} \]

where \(C\) is a normalization constant and \(P(\varepsilon; n)\) is the prior probability we assign to the efficiency before to consider the data. Therefore, since there is no reason to favour one value of the efficiency over another it makes sense to take:

\[ P(\varepsilon; n) = \begin{cases} 1 & \text{if } 0 \leq \varepsilon \leq 1 \\ 0 & \text{otherwise} \end{cases} \]

Finally, computing the normalization \(\int_{-\infty}^{+\infty} P(\varepsilon; k, n) \, d\varepsilon = 1\) and taking advantage of the Euler Beta function\(^2\), we can write explicitly the \(\varepsilon\) probability density function as:

\[ P(\varepsilon; k, n) = \frac{(n + 1)!}{k!(n - k)!} \varepsilon^k (1 - \varepsilon)^{n-k} \quad (5.2) \]

Now from the analytic form of the density function (Eq. 5.2) is possible to calculate the moments of the distribution:

\[ \bar{\varepsilon} = \int_0^1 \varepsilon P(\varepsilon; k, n) \, d\varepsilon = \frac{k + 1}{n + 2} \quad (5.3) \]

Moreover we can easily calculate also the mode of the distribution (most probable value) by solving \(dP/d\varepsilon = 0\), and we get:

\[ \text{mode}(\varepsilon) = \frac{k}{n} \quad (5.4) \]

---

\(^1\)Assuming that the sample space \(\Omega\) is divided among \(n\) mutually exclusive subset \(B_i\), \(\sum_{i=1}^{n} P(B_i) = 1\), if \(A\) is also a set belonging to \(\Omega\) the Bayes theorem states:

\[ P(B_i|A) = \frac{P(A|B_i) \cdot P(B_i)}{\sum_{j=1}^{n} P(A|B_j) \cdot P(B_j)} \]

\(^2\)\(B(\alpha + 1, \beta + 1) = \int_0^1 x^\alpha (1 - x)^\beta \, dx = \frac{\Gamma(\alpha + 1)\Gamma(\beta + 1)}{\Gamma(\alpha + \beta + 2)}\)

where \(\Gamma\) is the Euler Gamma function. For integer values holds: \(\Gamma(n + 1) = n!\).
We can easily see that now a good estimator for the efficiency \( \hat{\varepsilon} \) is the mode, because the mean for small value of \( n \) is biased. For this reason we make use of this estimator \( \hat{\varepsilon} = k/n \) for our computation of the detection efficiency, while the standard deviation is computed as \( \sigma_\varepsilon = \sqrt{V(\varepsilon)} \) where \( V(\varepsilon) \) is the variance of \( P(\varepsilon; k, n) \) (Eq. 5.5):

\[
V(\varepsilon) = \mu^2 - \varepsilon^2 = \frac{(k+1)(k+2)}{(n+2)(n+3)} - \frac{(k+1)^2}{(n+2)^2}
\] (5.5)

And now we can observe that \( \sigma_\varepsilon \) behaves correctly in the two limit cases of \( k = 0 \) and \( k = n \).

The detection efficiency for type 2 events (computed as described above for each bin in \( \xi \) ) is shown in figure 5.2. The black markers represent the efficiencies computed using Pythia6 generator, and the blue markers the ones computed for Pythia8. The two plots in the figure differ only for the binning and the range. Both of them display a narrow interval in \( \xi \) near to zero, which is the main region of interest for this study. The plot on the top shows the full rise of the efficiency that reaches fast the 90\% for a \( \xi \) of about \( 0.8 \div 0.9 \times 10^{-6} \) and the plateau value (really close to 1) for a \( \xi \) of about \( 3.5 \times 10^{-6} \). The plot on the bottom shows the efficiency in a smaller \( \xi \)-range and with a narrower binning with respect to the previous plot. This allows to highlight the different behaviour of the efficiency for the two simulations with a different generator. This discrepancy is more evident in the range \( 0.1 \times 10^{-6} < \xi < 0.5 \times 10^{-6} \) and is ascribable to the different modelling of the hadronic final state between Pythia6 and Pythia8.

Another important quantity that we want to compute with this study on SD processes, related to the detection efficiency, is the correction factor \( \varepsilon_{\text{tot}} \). This is dependent on the selected \( \xi \)-range where the measurement is performed and it has to take in account both the efficiency of detection and the “migration” ratio. In other word \( \varepsilon_{\text{tot}} \) has to take care of the fraction of events that are not detected by TOTEM inelastic detectors, due to their limited acceptance and efficiency, and of the fraction of events that are reconstructed (i.e. detected) but which came from outside of the considered \( \xi \)-range. This could happen because for a \( \xi_{\text{vis}} \) greater than
5.1 Study of Single Diffractive events

Figure 5.2: Efficiency of detection for inelastic SD events, as a function of $\xi$, for both Pythia6 (black marker) and Pythia8 (blue markers) generator. The plots on the top and the bottom differ only in binning and range, the second being a zoom of the first for small $\xi$ values.
the elastic limit \( (m_p^2/s) \) there is a non zero probability to detect also a process with a \( \xi \) lower than the one assumed as the cut value. In fact as the acceptance of the detectors is limited only toward small values of \( \xi \) there is no need to put an upper bound to its value, while there is the necessity of a lower bound. Once this bound is chosen (we will discuss on it further in the following text) the correction factor could be expressed as:

\[
\varepsilon_{\text{tot}} = \frac{1 - \varepsilon_{\text{migr}}}{\varepsilon_{\text{Reco}}}
\]

where: \( \varepsilon_{\text{migr}} = \frac{N^{\text{Reco}}_{\xi < \xi_{\text{vis}}}}{N^{\text{tot}}_{\xi < \xi_{\text{vis}}}} \) is the number of detected events for \( \xi < \xi_{\text{vis}} \) divided by the total number of reconstructed ones; \( \varepsilon_{\text{Reco}} = \frac{N^{\text{Reco}}_{\xi > \xi_{\text{vis}}}}{N^{\text{MC}}_{\xi > \xi_{\text{vis}}}} \) is the number of reconstructed events for \( \xi > \xi_{\text{vis}} \) divided by the number of MC “true” events with \( \xi > \xi_{\text{vis}} \).

Obviously \( N^{\text{tot}}_{\xi < \xi_{\text{vis}}} = N^{\text{Reco}}_{\xi < \xi_{\text{vis}}} + N^{\text{Reco}}_{\xi > \xi_{\text{vis}}} \) which means:

\[
\varepsilon_{\text{tot}} = \frac{N^{\text{MC}}_{\xi > \xi_{\text{vis}}}}{N^{\text{Reco}}_{\xi < \xi_{\text{vis}}}} \quad (5.6)
\]

We think that one possible logical choice for \( \xi_{\text{vis}} \) could be the value that minimizes at the same time the number of “lost” events and migration ones. Lost means that the process is not detected even if it has a \( \xi > \xi_{\text{vis}} \), while migration instead, as explained before, means that it is detected but it came from a \( \xi < \xi_{\text{vis}} \). This choice allows to quote a \( \xi \)-range, \( \xi > \xi_{\text{vis}} \), in which the detection is efficient at best without having a great contribution of migration. In other words, in this way we choose a region for which lost and migration events approach to balance each other. Two plots showing these quantities are in figure 5.3 in which the one on the top refers to Pythia6 generator and the other on the bottom to Pythia8. In these plots the black line shows per each bin, related to a particular value of \( \xi \), the number of migration events, while the blue line shows the number of lost ones. To make the plots clearer, both contributions have been divided by the total number of detected events. The red line instead shows the sum of the two contributions. Following this argument, it seems then clear that a reasonable choice for \( \xi_{\text{vis}} \) is between \( 0.21 \times 10^{-6} \) and \( 0.27 \times 10^{-6} \). In fact the minimum of the sum depends on the generator used for the simulation. Anyway the behaviour in both cases is similar and to us it
Figure 5.3: Percentage of lost events (blue line) and migration events (black line) with respect to the total number of detected ones, for various $\xi$ choices. The red line represents the sum of the two contributions. The plot on the top shows the results for SD processes simulation with Pythia6 generator and the plot on the bottom the ones for Pythia8 SD simulation.
Figure 5.4: Correction factor $\varepsilon_{\text{tot}}$ for various $\xi$ choices (left) and relative statistical error on it (right). For SD processes generated with both Pythia6 (top) and Pythia8 (bottom).

seems acceptable to choose a cut value of $\xi_{\text{vis}} \equiv 0.24 \times 10^{-6}$. This corresponds to a detection efficiency of about 40-60% depending on the generator employed for the simulation. However this could not be the only logical choice for the value of $\xi_{\text{vis}}$, but different thinking (and even simulation with other generators) could lead to prefer a different value. For these reasons the calculation of the correction factor, that relies on the equation 5.6, is performed on an extended $\xi$ interval. That is the usual one often used in this chapter for several plots. The results are reported in figure 5.4. The calculation of $\varepsilon_{\text{tot}}$ has been made for each value of $\xi_{\text{vis}}$ on various samples of MC events (50 samples of about 7500 events each, in this case). In the histograms on the left the mean value $\varepsilon_{\text{tot}}$ is reported for different choices of $\xi$ cut, while on the right the relative error on the correction factor is reported, due to the statistical uncertainty. The relative error has been calculated starting from the sample variance. The plots on the top of the figure refer to the simulation made with Pythia6 generator, while the ones on the bottom refer to Pythia8. Taking into
account the previous choice for $\xi_{vis} = 0.24 \times 10^{-6}$ we can see that the value of $\varepsilon_{tot}$ is $1.000 \pm 0.003_{(\text{stat})}$ for Pythia6 and $1.007 \pm 0.003_{(\text{stat})}$ for Pythia8. These two values, obtained with different generators, could also give us an idea of the systematics that the use of MC simulation induces in the evaluation of the correction factor for the inelastic cross-section measurement. From this simple argument it seems that the systematics is $\prec 1\%$, but further study on this, using other different generators (and even more statistics), is needed to be more conclusive about this point.

### 5.2 Study on Double Diffractive events

After Single Dissociative processes, the Double Dissociative ones are the greatest diffractive contribution to the total inelastic cross-section (see chapter 2, figure 2.5). Moreover also the DD detection efficiency, like the SD one, is $\xi$ dependant. Therefore, various samples of pure DD processes have been simulated starting from two different generators (Pythia6 and Pythia8 as usual) in order to repeat the study already done for Single Dissociative events. The $\xi$ distributions from the two kinds of simulation, made in the four different ways already mentioned in the previous section, are shown in figure 5.5. In which plot (a) shows the distribution for all Monte-Carlo DD generated processes. Plot (b) is made for events with at least one outgoing stable particle in the T1 or T2 detection range. Histograms (c) and (d) instead show the $\xi$ distribution for type 1 and type 2 selection criteria respectively. As for SD analysis: type 1 means that at least one hit is reconstructed in one of the two telescopes; type 2 that at least one track is detected. Focussing on the detection efficiency for these three different kinds of selection we can notice that for DD processes about the 95.8-95.5% of events are reconstructed by TOTEM inelastic detectors (relying on Pythia6 or Pythia8 respectively), with the request of one track. That is only $\sim 0.8$ percentage points less than requiring one hit, and means that about the $\sim 99.4\%$ of the processes with a stable particle in the T1-T2 range have at least one track reconstructed. Then also in the case of DD processes the type 2 criteria to select events detected from the TOTEM detectors seems a good choice (like for SD). Moreover referring to the previous chapter we can point out that, as
5.2 Study on Double Diffractive events

Figure 5.5: $\xi$ distribution of all the MC events (a); of the events with at least one stable particle in T1-T2 range (b); of type 1 events (c) and type 2 events (d). All the plots are made for Double Dissociative processes generated with Pythia6 (black line) and Pythia8 (blue starred line), all the curves are normalized to 1.

Expected, the DD processes are detected with a higher efficiency, respect to the SD ones. More precisely the TOTEM telescopes are about 14-15% more efficient for this kind of process.

Once the selection criteria has been chosen for Double Dissociation too, the next step is to study the dependence of the detection efficiency from $\xi$. For this purpose in figure 5.6 two plots showing this efficiency as a function of $\xi$ are reported. Both histograms has been computed using the Bayesian approach explained in the previous section, and are equal except for the range shown and the binning. The two colours, black and blue, indicate the different generator used in the simulation: black is used for Pythia6 and blue for Pythia8. The behaviour of these plots is similar to the ones in figure 5.2 that refers to SD processes. Anyway the rise seems
Figure 5.6: Efficiency of detection for inelastic DD events, as a function of $\xi$, for both Pythia6 (black marker) and Pythia8 (blue markers) generators. The plots on the top and the bottom differ only in binning and range.
faster, in fact the 90% is reached near $\xi \sim 0.6 \div 0.7 \times 10^{-6}$, while the plateau value (really close to 1) is reached at $\xi$ of about $2.5 \times 10^{-6}$. Moreover the efficiency from the two simulations (two different generators) still behave differently, in particular in the range $0.12 \times 10^{-6} < \xi < 0.5 \times 10^{-6}$.

Following the same arguments pointed out in the previous section for SD processes, also for DD events it’s important to verify if the $\xi_{vis}$ decided before as a cut value (in order to restrict the inelastic cross-section measurement to a sub-range $\xi > \xi_{vis}$) is still a good choice, and then to calculate the appropriate factor $\epsilon_{dd tot}$ to correct the number of measured DD processes in order to obtain the “true” inelastic rate. As for the previous study the $\epsilon_{dd tot}$ has to take into account the lost events (due to detector limited acceptance and efficiency) and of migration ones (reconstructed event coming from outside the $\xi$-range). For these reasons it has been calculated in a completely similar way (see equation 5.6). To check the choice of $\xi_{vis}$, in figure 5.7 are reported two plots made for DD processes, that are equivalent to the ones of figure 5.3. These two plots (figure 5.7) show per each bin (that refers to a certain $\xi$ value): the number of migration events (black line), the number of lost events (blue line) and the sum of the two contributions (red line). Each of the three histograms is then scaled over the total number of detected processes, to make the plots more readable. The behaviour of these histograms is quite similar to the one shown in figure 5.3, even if the percentage contribution of lost and migration events is reduced. This is expected, because we handle Double Dissociation like two SD processes out of which we select the one with the higher $\xi$, and for this reason there are less DD events than SD ones at very low $\xi$ value. This is shown in figure 5.8 in which the $\xi$ distribution (for all MC events) are represented, both for SD processes (black line) that for DD ones (red starred line). To compare the two distributions even if they have different statistics they are both scaled on the respective number of entries. Only the plot referring to Pythia8 generator is shown, but the one for Pythia6 has a really similar behaviour.

The minimum of the sums of the two contribution is anyway between about $0.18 \times 10^{-6}$ and $0.25 \times 10^{-6}$ in $\xi$ (depending on generator), so the choice of $\xi_{vis} = 0.24 \times 10^{-6}$, previously made looking only at Single Diffraction, still seems reasonable.
Figure 5.7: Percentage of lost events (blue line) and migration events (black line) with respect to the total number of detected ones, for various $\xi$ choices. The red line represents the sum of the two contributions. The processes analysed are Double Dissociative and are generated with Pythia6 (top) and Pythia8 (bottom).
Moreover this implies that the detection efficiency for DD processes is always higher than 40-60\% in the range of $\xi > 0.24 \times 10^{-6}$.

Achieved this important result, what remains to do is the computation of the correction factor $\varepsilon_{dd}$, that is performed not only for the chosen value of $\xi_{vis}$, but for different values in the range $0 < \xi_{vis} < 1 \times 10^{-6}$. The calculation is made in the same way described for SD and the results are reported in figure 5.9. The only difference is in the statistics generated with Pythia8, that consists in 40 samples of about 5000 events each, instead of the usual 50 samples of about 7500 events (this choice is due only to the long simulation time). The plots on the left of the figure show the $\varepsilon_{dd}$ for different choices of $\xi_{vis}$, while the ones on the right show the relative error on the correction factor, due to the statistical uncertainty. The plots on the top of the figure display the results for the simulation made with Pythia6 generator, then the others on the bottom refer to Pythia8. Looking at the decided cut value $\xi_{vis} = 0.24 \times 10^{-6}$ we can see that it corresponds to a $\varepsilon_{dd}$ of $1.001 \pm 0.002_{\text{stat}}$ for Pythia6 and $1.003 \pm 0.002_{\text{stat}}$ for Pythia8. So also for the DD events it seems that the systematics is $< 1\%$; however, also in this case, further study using other different generators (and even more statistics) is needed to be more conclusive.
Figure 5.9: Correction factor $\varepsilon_{\text{dd}}$ for various $\xi$ choices (left) and relative statistical error on it (right). For Double Dissociative processes generated with Pythia6 (top) or Pythia8 (bottom).
5.3 A brief looking upon Non Dissociative events

The greatest contribution to the total inelastic cross-section is represented by Non Dissociative processes. For this reason, even if the detection of these events is less problematic, it is important to perform a brief study also on these processes. In order to understand the detection efficiency and other issues related to them. In particular we want to understand if also for these processes it is needed a correction factor to evaluate the number of “true” ND events contributing at the inelastic cross-section. For this purpose, as usually, we used two different simulations, always for ND processes only, with the two different generators Pythia6 and Pythia8. Then, believing that the choice of the type 2 criterion could be still appropriate to define an event as detected or not, we start checking the detection efficiency for this kind of selection criteria only. This is done simply plotting for each simulated event the number of tracks in the T1 telescopes versus the number of tracks in T2. As is depicted in figure 5.10, in which the 2D histogram on the left is made with the simulation relying on Pythia6, while the one on the right is computed using the Pythia8 based simulation. In both cases we can notice that the reconstruction efficiency is practically 100%. In fact for Pythia8 simulation only 1 event, over the 109910 generated ones, has no reconstructed tracks in both the inelastic telescopes and, according with the chosen criterion, it’s to consider as not detected. While for Pythia6 only 3 events over 175911 do not pass the selection cuts. So also for the ND processes, even more than for SD and DD ones, the decision to require one track

![Figure 5.10: 2D histograms showing the number of tracks in T1 telescopes versus the number of tracks in T2. Using Pythia6 (left) and Pythia8 (right) generators.](image-url)
Figure 5.11: $\xi$ distribution of all ND processes with a rapidity gap greater than 3 $\eta$ unit. For two simulations made with Pythia6 (left) and Pythia8 (right) generators.

as detection criterion seems good, and there is no need to investigate over different selection cuts. Moreover there is no need of a Non Dissociative correction factor to balance the limited detection efficiency. Anyway an interesting thing to study could be the ratio of ND events that “imitate” the DD ones. For example if in a future common data taken with CMS one would want to distinguish between the different diffractive contributions using topology tagging. From this prospective we show in figure 5.11 the $\xi$ distribution (in a narrow interval close to zero) of the ND events that have a rapidity gap of at least 3 unit in $\eta$, ($\Delta \eta > 3$ is a used cut to select DD events [28]). For these events $\xi$ is computed in the usual way described in the first part of this chapter. The plot on the left is made using the Pythia6 generator for the simulation, while the plot on the right is made using Pythia8. Both of them show that the ratio of ND events passing this requirement is about 3-4% (depending on the generator). It is quite small but probably if one thinks to separate the different contribution needs to take in account this contamination of ND to DD processes. Moreover none of these events have a $\xi$ smaller than the decided cut value $\xi_{vis} = 0.24 \times 10^{-6}$.

This concludes our study on inelastic processes, but since the contribution to the inelastic cross-section of the remaining diffractive events (Double pomeron exchange, Multi pomeron exchange, etc.) is much smaller, we can argue that the the choice made for $\xi_{vis} = 0.24 \times 10^{-6}$ is a good one. So the TOTEM experiment, due to its good coverage in a high rapidity region, has good prospective to perform the measurement
of the inelastic cross-section in the kinematic range $\xi > 0.24 \times 10^{-6}$, corresponding to a diffractive mass $M_x > 3.4$ GeV at C.M. energy of 7 TeV. Moreover relying on Pythia8 the correction factors for SD and DD events are respectively $\varepsilon_{tot} = 1.007 \pm 0.003_{(stat)} \pm 0.007_{(syst)}$ and $\varepsilon_{tot} = 1.003 \pm 0.002_{(stat)} \pm 0.003_{(syst)}$, where the systematic error takes in account only the uncertainty due to the fragmentation model and is computed from the difference between Pythia6 and Pythia8 prediction. To take in account the systematic arising from the uncertainty in the underlying $\xi$ distribution a simulation with a generator (e.g. Phojet) that uses a different model (for the dependency of the diffractive cross-section on $\xi$) is needed. Moreover the other more relevant issues for the measurement of the inelastic cross-section, not discussed here, are: the background contribution (especially beam-related background) and the trigger efficiency.
5.3 A brief looking upon Non Dissociative events
Conclusions

The TOTEM experiment at the CERN LHC has been designed for TOTal cross-section, Elastic scattering and diffractive dissociation Measurements. Both total cross-section and diffractive dissociation measurements require a good understanding of the TOTEM inelastic detectors T1 and T2. This thesis work has been done in this perspective, during which the candidate has performed a tuning of the simulation for the T2 telescope, a study of the noise contribution on the same detector and a preliminary study on the inelastic processes with both the telescopes T1 and T2.

The work on the simulation tuning was needed due to the big mismatch found between the data and the simulation itself, as observed in a first study of the T2 detector occupancy. The tuning consisted of two parts mainly: a first one where the optimisation of some cut parameters is achieved (in order to have a comprehensive and performing simulation); and a second one in which the simulated geometry is checked and corrected. We have finally reached a stable setting for the simulation, in which the CPU time usage has been optimised and the mismatch with respect to data is not anymore related to a wrong use of the Geant4 cuts setting. Moreover, all volumes of the CMS detector of interest for the T2 response are now included in the simulation and the geometry of the T2 region has been optimised. Overall, looking back to the starting point of our work, the tuning of the simulation led to appreciable results, meaning that the simulation is now much closer to the data. Anyway there is still a non negligible residual difference between observed data and simulation. The possible source of this discrepancy could be searched in problems related to the propagation of the simulated particles through the different volumes. Unfortunately at the time we have not the possibility to act directly on this problem and verify this issue, but a solution for this problem is known and is ongoing.
At least in principle another source for this discrepancy could be the noise contribution in data. In fact noise is actually not simulated a priori (with the only exception of the capacitive pad and strip noise). Therefore it was important to determine the necessity of a simulation improvement in this sense. This motivated the work of the candidate to understand the noise incidence and features for the T2 detector. After this comprehensive study reported in the fourth chapter of this thesis, we can conclude that the noise is not a relevant issue for the T2 detector, and there is no need to simulate it. In particular we can affirm that in situations of very low occupancy the detector is really silent. Anyway also for what concerns the noise induced by the passage of charged particles, this is more difficult to study and evaluate, but it seems anyway under control.

After this long and detailed study to optimise the detector simulation and to understand noise related features, we finally moved to simulate inelastic proton-proton collisions and study the detectors response.

In particular we would like to verify in more detail the acceptance of the detectors to inelastic events in order to point out the perspective for an inelastic cross-section measurement with the TOTEM detectors. During this study we analysed the response for Single Dissociative and Double Dissociative diffractive processes and for the Non Dissociative processes, that together constitute the greatest part of the inelastic cross section. In particular this study points out the really good efficiency of the TOTEM telescopes T1 and T2 in detecting inelastic processes down to a very small value of the proton fractional momentum loss $\xi$. And finally, thanks to this, we can affirm that the TOTEM experiment, due to its good coverage in a high rapidity region, has very good perspectives to perform the measurement of the inelastic cross-section in a wide kinematic range for $\xi > 0.24 \times 10^{-6}$. 
Acknowledgements

At the end of this thesis work, it is a pleasure for me to thank all the people that made this thesis possible. In particular I would like to express my gratitude to Dr. Stefano Lami and Dr. Giuseppe Latino, for their tireless help, support and teaching. Thanks for your constant assistance, guide and patience. I’m pleased to thank Mirko Berretti and Dr. Eraldo Oliveri for their uninterrupted aid to my thesis and to me with their knowledge and friendship. I am grateful to all the TOTEM people, at CERN and not, that helped me during this period. I’m also indebted with all my colleagues and friends at the university and at CERN for providing a stimulating and fun environment in which to learn and work. Moreover a special thanks goes to all my friends that stand by me since a long time. Thanks Guys. Lastly, and most importantly, I wish to thank my parents and my family, for all the great and small things they do everyday for me. To them I dedicate this thesis.

Alla fine di questo lavoro di tesi, vorrei spendere due parole per ringraziare tutti coloro che mi hanno aiutato e ne hanno reso possibile la realizzazione. Prima di tutto voglio ringraziare nella maniera più sincera e sentita il Dott. Stefano Lami ed il Dott. Giuseppe Latino che durante questo lavoro di tesi sono sempre stati pronti ad aiutarmi con pazienza e dedizione, guidandomi ed insegnandomi molto. Un ringraziamento particolarmente grande va poi a Mirko Berretti ed Eraldo Oliveri che durante questo lungo periodo mi hanno aiutato molto, sia nella parte tecnica, con la loro conoscenza, che in quella umana, con la loro amicizia. Non posso quindi dimenticare di ringraziare tutte le persone del gruppo di TOTEM che mi hanno aiutato in questo periodo sia al CERN che non. Inoltre voglio ringraziare tutti i
miei amici e colleghi dell’università di Pisa, di Siena e del CERN per il loro supporto nello studio come nel tempo libero. Ai miei amici di sempre dedico un grazie speciale, perché da molto tempo condividono con me i momenti importanti rendendoli più veri, ed il tempo che spendo con loro mi rinfranca sempre. Infine il ringraziamento più importante va alla mia famiglia, insieme alla dedica di questo lavoro. Perché sono loro che con piccoli o grandi gesti mi sostengono ogni giorno, dando senso a questo come agli altri momenti della vita.
Bibliography


C. Bourrely et al., Impact-picture phenomenology for $\pi^\pm p$, $K^\pm p$ and pp, $\bar{p}p$ elastic scattering at high energies, Eur. Phys. J. C28, 97 (2003).


