Università degli Studi di Siena

Facoltà di Scienze Matematiche Fisiche e Naturali

Corso di Laurea in Fisica Sperimentale

Tesi di Laurea Specialistica

Development of reconstruction algorithms for inelastic processes studies in the TOTEM experiment at LHC

Relatore: Prof. G. Latino
Candidato: M. Berretti

Correlatore: Prof. A. Brez

Anno Accademico 2007/2008
Sessione Estiva - 4 Luglio 2008
Contents

Introduction .......................................................... 5

1 The TOTEM experiment at the LHC ............................. 7
   1.1 The LHC machine ............................................. 7
      1.1.1 General aspects ....................................... 7
      1.1.2 Beam optics ............................................ 11
   1.2 Overview on TOTEM detectors ............................... 15
      1.2.1 Roman Pots ............................................ 17
      1.2.2 T1 telescope .......................................... 19
      1.2.3 T2 telescope .......................................... 20
   1.3 The beam pipe ................................................. 24

2 TOTEM Physics programme ...................................... 29
   2.1 Overview on TOTEM Physics programme .................... 29
   2.2 Total cross section measurement ........................... 30
      2.2.1 Total cross section at LHC ........................... 30
      2.2.2 Measurement method ................................... 32
   2.3 Elastic cross section measurement ......................... 35
   2.4 Diffractive and inelastic processes ......................... 37
      2.4.1 Diffractive processes ................................ 37
      2.4.2 Proton diffraction into three jets ................... 39
      2.4.3 Tuning of cosmic ray Monte Carlo generators ... 41
   2.5 Low x physics ................................................. 43
## CONTENTS

2.5.1 PDF saturation .................................................. 44
2.5.2 Minijet .............................................................. 44
2.5.3 Forward Drell-Yan ............................................... 46
2.5.4 Mueller Navalet jets ........................................... 48

3 Track reconstruction with the T2 detector .......................... 51
3.1 Simulation and reconstruction chain ................................. 51
3.2 Hit reconstruction ................................................... 56
3.3 Road finder ............................................................ 59
3.4 Track reconstruction .................................................. 63
3.5 Tracking performance ................................................. 66
3.5.1 Beam pipe effect ................................................... 66
3.5.2 Primary track selection ......................................... 69
3.5.3 Angular resolution on track reconstruction ................... 77
3.5.4 Reconstruction of tracks from neutral pion decay .......... 80

4 Topological jet algorithms ............................................. 85
4.1 Jets and jet algorithms .............................................. 86
4.2 Monte Carlo samples .................................................. 88
4.2.1 Di-jet events ....................................................... 89
4.2.2 Single Diffractive events ..................................... 90
4.3 Topological jet algorithms .......................................... 90
4.3.1 Cone algorithm .................................................... 90
4.3.2 $k_t$ algorithm ..................................................... 95
4.3.3 Jet parameters .................................................... 98
4.4 Algorithm setting .................................................... 102
4.4.1 Algorithm parameters .......................................... 102
4.4.2 Jet selection cuts ................................................ 105
4.5 Results and comparisons ........................................... 106
4.5.1 Performance at particle level .................................. 107
4.5.2 Performance at track level ..................................... 114
CONTENTS

Conclusions 121

Acknowledgment 125

Bibliography 127
Introduction

The TOTEM experiment [1] at the Large Hadron Collider (LHC) is designed and optimized to measure the total pp cross section at a center of mass energy of $E = 14$ TeV with a precision of about $1\pm2\%$, to study the nuclear elastic pp cross section over a wide range of the squared four-momentum transfer ($10^{-3} \text{GeV}^2 < |t| < 10 \text{GeV}^2$) and to perform a comprehensive physics program on diffractive dissociation processes, partially in cooperation with the CMS experiment. In order to fulfill this physics programme the TOTEM experiment has to cope the challenge of triggering and recording events in the very forward region with a good acceptance for particles produced at very small angles with respect to the beam. Based on the “luminosity independent method”, the evaluation of the total cross section with such a small error will in particular require simultaneous measurement of the pp elastic scattering cross section $d\sigma/dt$ down to $|t| \sim 10^{-3} \text{GeV}^2$ (to be extrapolated to $t = 0$) as well as of the pp inelastic interaction rate. In particular, the detection of elastically scattered protons at a location very close to the beam is required together with particle detection with the largest possible coverage in order to reduce losses on inelastic events detection to a few percent.

The TOTEM physics programme will be accomplished by using three different types of detectors: elastically scattered protons will be detected by Roman Pots detectors (based on silicon technology) placed about at 147 m and 220 m from the interaction point; inelastic processes will be detected by two tracking telescopes, T1 and T2 (based on gas detectors), embedded into the forward region of the CMS experiment.

This thesis is structured in four chapters. Chapter 1 begins with an introduction to the LHC and to the beam dynamics. The TOTEM apparatus is then described with particular emphasis given to the T2 detector, on which the work reported in this thesis is related.
This chapter also includes the description of the beam pipe around the interaction point, which is of interest for track reconstruction with TOTEM inelastic telescopes because of the possibility of multiple scattering and secondary particle production. In chapter 2 an overview on the TOTEM physics program is given and a description of some possible processes involving measurement with the T2 detector is reported. In chapter 3 the thesis work related to the development of algorithms for geometrical hit and track reconstruction in T2 is presented. The former reconstructs the geometrical position of the charged particle which ionize the active zone of the T2 triple-GEM chambers, starting from the digital signal of the activated read-out channels; the latter fits the reconstructed hits in order to obtain the particle track through the telescope. Moreover, in this chapter some studies for the optimization of the track selection criteria are reported, which in particular will be important for primary vertex reconstruction in inelastic events and hence for the measurement of the total cross section. In chapter 4 it is then investigated the possibility of hadronic jet reconstruction with the TOTEM inelastic telescopes, only relying on topological information $(\eta, \phi)$ of charged particles. For this purpose, two innovative jet algorithms have been developed and tested at the particle level and then also applied at the reconstructed track level on Pythia Di-jet and Single Diffractive events, which allow jet reconstruction with a reasonable efficiency and relatively low fake-jet rate. These algorithms can in principle be combined with the energy information from the CMS CASTOR calorimeter for an optimal jet reconstruction in the very forward region, to be used in several analyses of interest.
Chapter 1

The TOTEM experiment at the LHC

1.1 The LHC machine

The Large Hadron Collider (LHC) is a superconductor synchrotron which has been installed at CERN (European Organization for Nuclear Research) and which start-up is scheduled for summer 2008. In this section a general overview on the LHC technology, performance and potential issues during its operation is firstly given. Subsequently, the typical machine optical parameters, used to describe the beams dynamics, are introduced.

1.1.1 General aspects

The LHC is installed in the 27 Km tunnel where the CERN Large Electron Positron (LEP) collider was placed before. When fully operative, it will be the most powerful hadron collider ever built, allowing proton-proton collisions up to a peak luminosity of $10^{34} \text{cm}^{-2} \text{s}^{-1}$ at a center of mass (C.M.) energy of 14 TeV (respectively about 100 times and 10 times the highest luminosity and collision C.M. energy ever reached).

The physics studies allowed by the LHC span most of the biggest questions in high energy frontier physics, from production/detection of supersymmetric particles to the confirmation or exclusion of the standard model Higgs field. However, entering into a new energy scale, something new in our knowledge can come from unexpected physics.
The LHC will accelerate two separate beams of protons\(^1\) (each composed, at the peak luminosity, by 2808 bunches of about \(10^{11}\) protons) up to an energy of 7 TeV, and then will bring them into head-on collisions at four interaction points (IP) where different detector systems (ALICE, ATLAS/LHCf, CMS/TOTEM, LHCb) are located.

The acceleration of the protons up to the final 7 TeV energy is not exclusively done by the LHC. The first stage of the acceleration is indeed done by the CERN accelerator complex (see fig. 1.1), which is a succession of particle accelerators where the protons reach increasingly higher energies. The protons obtained from dissociation of hydrogen gas are first accelerated up to 50 MeV in the LINAC2 linear accelerator. They are then sent into the PS Booster, where they reach an energy of 1.8 GeV, and then they are accelerated up to 26 GeV inside the Proton Synchrotron (PS). The last pre-acceleration stage, up to 450 GeV, is done by the Super Proton Synchrotron (SPS). After that the protons are injected in the Large Hadron Collider (LHC), which will bring them to the final 7 TeV energy.

![Figure 1.1: The CERN accelerator complex.](image)

As it will accelerate two counter-rotating particle beams of the same charge, the LHC consists of two “superconducting magnetic channels” or “rings” housed in the same yoke

\(^1\)In a second stage the LHC will also collide beams of heavy ions, such as lead, with a collision energy up to 1148 TeV.
1.1 The LHC machine

and cryostat (see fig. 1.2), a unique configuration that not only saves space but also gives a 25% cost saving with respect to separate rings.

**LHC DIPOLE : STANDARD CROSS-SECTION**

Figure 1.2: LHC cryogenic structure for two 8.36 T dipole magnets.

In order to bend 7 TeV protons around the ring the LHC dipoles must be able to produce magnetic fields of 8.36 Tesla, a value which can be reached only by using superconducting magnets. The LHC magnetic system consists of 1296 superconducting dipoles (which are cooled down to 1.9 K by using superfluid helium, see fig. 1.2) and more than 2500 other magnets, which allow to guide and focus the beams toward the interaction points around which the experiments are built. Table 1.1 summarizes some of the main LHC machine parameters at the peak luminosity [2] (the last two parameters will be explained in next section).
1.1 The LHC machine

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beam energy</td>
<td>7.0 TeV</td>
</tr>
<tr>
<td>Dipole field</td>
<td>8.3 T</td>
</tr>
<tr>
<td>Luminosity</td>
<td>$10^{34}$ cm$^{-2}$ s$^{-1}$</td>
</tr>
<tr>
<td>Bunch spacing</td>
<td>24.95 ns</td>
</tr>
<tr>
<td>Particles per bunch</td>
<td>$10^{11}$</td>
</tr>
<tr>
<td>Beam lifetime</td>
<td>22 h</td>
</tr>
<tr>
<td>Luminosity lifetime</td>
<td>10 h</td>
</tr>
<tr>
<td>Normalized transverse emittance</td>
<td>3.75 $\mu m$ rad</td>
</tr>
<tr>
<td>Beta values at IP 5</td>
<td>0.55 m</td>
</tr>
</tbody>
</table>

Table 1.1: LHC main machine parameters at the peak luminosity [2].

The machine instantaneous luminosity $\mathcal{L}$ is defined according to [3]:

$$\mathcal{L} = \frac{f N_1 N_2}{4\pi \sigma_x^* \sigma_y^*}$$

(1.1)

where $\sigma_x^*$ and $\sigma_y^*$ are the transverse gaussian beam profiles in the horizontal (bend) and vertical direction at IP, $N_1$ and $N_2$ are the number of protons in the colliding bunches and $f$ is the crossing frequency ($f = f' \cdot n$, where $f'$ is the single bunch revolution frequency and $n$ is the number of bunches in a beam). A high luminosity can so be achieved by using bunches having small cross section at the IP and/or high bunch crossing frequency and/or highly populated bunches. However, such requirements make the machine sensitive to several problems, some of which are reported below.

A high bunch focusing can lead to the “beam-beam effect” : when two bunches cross at the interaction point, only a tiny fraction of particles collide head-on to produce events from hadronic interaction; all the others are deflected by the strong electromagnetic field of the opposing bunch and this deflection is stronger for denser (smaller area) bunches; these deflections accumulate turn after turn and may eventually lead to particle losses. Experience from previous colliders showed that the bunch density cannot be increased beyond a certain limit in order to preserve a sufficiently long beam lifetime. Therefore, in order to reach the desired luminosity, the LHC has to operate as close as possible to this limit. Anyway, a fraction of particles will diffuse towards the beam pipe wall and will be lost. In this cases the particle energy is converted into heat in the surrounding material which can “quench” a magnet out of its cold, superconducting state. When it occurs,

$^2$The resistive transition from the superconducting to the normal-conducting state is called quench.
Unless precautions are taken, the stored magnetic energy may cause severe damages to the magnet. Therefore, a reliable active quench protection circuit is needed to bring safely the current down to zero when a quench occurs. Moreover, a collimation system is installed in order to catch the unstable particles before they can reach the beam pipe wall, so as to confine losses in well shielded regions away from any superconducting element.

The number of particles in each bunch cannot be increased indefinitely because the event pile-up becomes more and more important when this number increases. For instance, at the peak luminosity configuration, about 20 events of soft hadron interactions are expected for each beam crossing, which have to be disentangled from a more interesting hard interaction (TOTEM will indeed have a reduced possibility of measurement in this scenario). Furthermore, the bunch crossing rate cannot be too high, being limited by the current development of suitable fast-response detectors and read-out systems to be used in the experiments.

1.1.2 Beam optics

In this section some typical quantities which are used in the description of the dynamics of the circulating beams are introduced. In particular, the transverse motion of a particle in the xy plane (x is the horizontal bend axis, y is the vertical axis and z is the beam axis) is described using the conjugate coordinates \( x(s), x'(s) \equiv \frac{dx}{ds} \) and \( y(s), y'(s) \equiv \frac{dy}{ds} \) where \( s \) can be thought as the coordinate on an ideal circular orbit. It can be noted that particles which circulate almost parallel to the nominal beam axis have \( x'(s) \equiv \frac{dx}{ds} \equiv \tan \theta_x \sim \theta_x \) and \( y'(s) \sim \theta_y \). The meaningful informations of the beam dynamics are given in terms of collective quantities which represent all the trajectories of the particles. For instance, as described later, the envelope of all such trajectories is described by the betatron function \( \beta(s) \), which is determined by the accelerator magnet configuration.

The motion of a charged particle in a lattice consisting of only dipoles and quadrupoles is described, in terms of the conjugate coordinates already introduced, by the Hill equations [4]. For the y coordinate:

\[
y'' - k(s)y = 0 \tag{1.2}
\]

This equation looks like the one of a harmonic oscillator with a s-dependent oscillation
1.1 The LHC machine

frequency. The equation for the x coordinate is similar, but has another term in $x(s)$ which takes into account that the motion is constrained in a circular trajectory. The solution of the above equation is usually wrote as:

$$y(s) = \sqrt{\varepsilon} \sqrt{\beta_y(s)} \cos(\psi(s) + \phi)$$  \hspace{1cm} (1.3)

where the transverse emittance $\varepsilon$ and the betatron function $\beta_y(s)$ (for the y coordinate) are introduced\(^3\). With the introduction of two convenient functions, $\alpha(s)$ and $\gamma(s)$ which are a combination of $\beta(s)$ and its derivatives, it is possible to show that the above solution $y(s)$ and its first derivative $y'(s)$ describe an ellipse having its shape and orientation which depend on $s$, but with constant area $\pi \varepsilon$ throughout the ring\(^4\). This is shown in fig. 1.3 (from ref. [4]) where the maximum amplitude of the beam in the y coordinate (i.e. the transverse beam size in y) is $y_M(s) = \sqrt{\varepsilon \beta(s)}$. It is also clear that, whenever the ellipse axes are parallel to $y$ and $y'$, the maximum angular divergence is $y'_M(s) = \sqrt{\varepsilon / \beta(s)}$. In this case the transverse beam size and the angular divergence in the y coordinate at a given point $s$ in the storage ring can be written as:

$$\sigma_y(s) = \sqrt{\varepsilon \beta(s)}$$  \hspace{1cm} (1.4)

and

$$\sigma_{\theta_y}(s) = \sqrt{\varepsilon / \beta(s)}$$  \hspace{1cm} (1.5)

The formalism introduced above makes it also possible to write in matrix form the evolution of the conjugate variables $(y(s), y'(s))$ from a point $s = s_0$ to $s = s_1$. The transfer matrix, which links $(y(s_0), y'(s_0))$ to $(y(s_1), y'(s_1))$, is known for each configuration of the magnet system and it is possible to write its matrix elements in terms of the accelerator functions introduced above. For instance, if $s$ is the position of a TOTEM Roman Pot (RP) detector, $(y(s), y'(s))$ are the y conjugate coordinates of a proton in the RP and $(y^*, y'^*)$\(^5\) are the proton conjugate coordinates at the interaction point, the relation among the

\(^3\)Hereafter, the notation is simplified by using $\beta(s)$ instead of $\beta_y(s)$ or $\beta_x(s)$. The coordinate of interest will be clear from the context.

\(^4\)This is a direct consequence of the Liouville theorem, which is considered valid because it is assumed that the particle energy is a constant as beam-beam scattering and synchrotron radiation are neglected.

\(^5\)A generic variable $X$, when related to IP, is usually quoted as $X^*$. 
1.1 The LHC machine

Figure 1.3: Phase space ellipse describing the \((y, y')\) particle coordinates; \(\gamma\) and \(\alpha\) are function of \(\beta(s)\) and its derivatives \([4]\).

The properties of the beam optics (i.e. the elements of \(M(s)\)) can be expressed by the two optical functions \(v(s) = M_{11}(s)\) (magnification) and \(L(s) = M_{12}(s)\) (effective length) which, at a distance \(s\) from the IP, are defined by the betatron function \(\beta(s)\) and the phase advance \(\Delta \mu(s)\) according to:

\[
\begin{pmatrix}
    y(s) \\
    y'(s)
\end{pmatrix}
= \begin{pmatrix}
    \gamma \\
    \alpha
\end{pmatrix}
= M(s)
\begin{pmatrix}
    y^* \\
    y'^*
\end{pmatrix}
\]  \hspace{1cm} (1.6)

\[
v(s) = \sqrt{\frac{\beta(s)}{\beta^*}} \cos \Delta \mu(s)
\]

\[
L(s) = \sqrt{\beta(s)\beta^*} \sin \Delta \mu(s)
\]

with \(\Delta \mu(s) = \int_0^s \frac{1}{\beta(s')} ds'\)  \hspace{1cm} (1.7)

Where \(\beta^*\) is the betatron function at the IP. The transverse y-displacement \(y(s)\) of the
1.1 The LHC machine

proton in the Roman Pot is thus expressed by the equation:

\[ y(s) = v_y(s) \cdot y^* + L_y(s) \cdot y'^* \]  \hspace{1cm} (1.8)

From the above equations we can see that, in order to eliminate the dependence on the transverse position of the proton at the collision point, the magnification \( v(s) \) has to be chosen as close to zero as possible (parallel-to-point focusing, \( \Delta \mu(s) = \pi/2 \)). It is also clear that the larger is the effective length (having fixed the minimum \( y(s) \) which can be reached by the RP active zone), the smaller is the scattering angle at IP which is possible to measure. Equations 1.4 and 1.5 also show that configuring the magnets in order to have a high \( \beta^* \) leads to a large beam size as well as to a small angular divergence at the IP, the latter being required for a precise measurement of small scattering angles for the protons.

Similar expressions hold for particle displacement in the x direction, with the only important difference related to the fact that the x coordinate, as a consequence of the bending magnetic fields, is also sensitive to an eventual proton energy-loss. Defining \( \xi \equiv \Delta p/p \) as the relative energy loss of the proton, the detected \( x(s) \) position at the RP location is given by:

\[ x(s) = v_x(s) \cdot x^* + L_x(s) \cdot x'^* + \xi \cdot D(s) \]  \hspace{1cm} (1.9)

where \( D(s) \) is dispersion of the machine. Most of diffractive processes are characterized by the presence of at least one proton which lost a fraction \( \xi \) of its energy during the interaction. Assuming \( v_x(s) = 0 \) (parallel-to-point-focusing), the previous equation shows that the measurement of the momentum loss \( \xi \) with an acceptable resolution imposes \( L_x(s) \sim 0 \) in order to eliminate the dependence of \( \xi \) on \( \theta^*_x \), the horizontal scattering angle at the IP.
1.2 Overview on TOTEM detectors

The TOTEM experiment is composed by three different detectors located symmetrically on both sides of the interaction point IP5, the same LHC experimental area as CMS [5] (see fig 1.4). As TOTEM is not a general purpose experiment, its detectors cover only a particular acceptance region. Its physics programme requires in fact a good acceptance for particles produced at angles close to the beam axis. The combination of the CMS and TOTEM experiments represents however the largest acceptance detector ever built at a hadron collider which will allow, as complementary to the total and elastic cross section measurement made by TOTEM, the study of a wide range of physics processes in diffractive as well as non-diffractive interactions with an unprecedented coverage in pseudo-rapidity\(^6\).

For this purpose the TOTEM data acquisition system (DAQ) is designed to be compatible with the CMS DAQ in order to make common triggering and data taking possible at a later stage.

The TOTEM experimental setup comprises Roman Pots detectors, to measure leading protons elastically scattered at very small angles within the beam pipe, and the T1 and T2 inelastic telescopes, providing charged track reconstruction for \(3.1 < |\eta| < 6.5\) with a \(2\pi\) coverage and with a very good efficiency in order to minimize losses (see fig.1.4, top). As energy flow and charged particle multiplicity in inelastic events peak in the forward region (see fig. 1.5), the location of TOTEM inelastic telescopes will ensure the detection of about 99.5\% of all non-diffractive minimum bias events and of about 84\% of all diffractive events. Hence, almost all inelastic events (\(\sim 95\%\)), having charged particles within the geometrical acceptance of T1 or T2, will be triggerable with these detectors. The inelastic telescopes will also provide the reconstruction of the primary vertex, which will allow to reject background events (mainly beam-gas interactions and halo muons). It is important to remark that, as all TOTEM detectors are outside the strong CMS magnetic field (central) region, the track reconstruction of charged particles they provide does not give information on the particle momentum. However, as the T1 and T2 telescopes will provide tracking in front of the CMS HF (T1) and Castor (T2) very forward calorimeters, the combination of these detectors will allow, for instance, a more

\(^{6}\)The particle pseudo-rapidity \(\eta\) is defined as \(\eta = -\ln(\tan \frac{\theta}{2})\), where \(\theta\) is the polar scattering angle of the particle with respect to the original beam direction.
1.2 Overview on TOTEM detectors

Figure 1.4: The TOTEM detectors placed around IP5. Top: the TOTEM forward trackers T1 and T2 embedded into the forward region of the CMS detector. Bottom: TOTEM Roman Pots location along the LHC beam line at a distance about 147 m (RP147) and 220 m (RP220) from the interaction point IP5, RP180 being another possible location at the moment not equipped. All TOTEM detector components are located on both sides of IP5.

A complete study of diffractive processes, low-x phenomena and particle/energy flows in the very forward region (see chapter 2).

The read-out of all TOTEM detectors is based on the digital VFAT chip [6], which is a tracking front-end ASIC specifically designed for the TOTEM experiment and characterized by trigger capabilities.

In this section a general overview on TOTEM detectors is given. In particular, a more detailed description is dedicated to the T2 telescope (sub-section 1.2.3), since this thesis is focused on the development of reconstruction algorithms for signals generated by this detector.
1.2 Overview on TOTEM detectors

1.2.1 Roman Pots

The detection of very forward protons is performed by movable beam insertions called “Roman Pots” (RP). The Roman Pot is an experimental technique introduced at the ISR for the detection of forward protons in elastic or diffractive scattering events. It has been successfully employed in other colliders like the SPS, TEVATRON, RHIC and DESY. Silicon detectors are placed inside a secondary vacuum vessel, called “pot”, and moved into the primary vacuum of the machine through vacuum bellows. In this way the detectors are physically separated from the primary vacuum which is so preserved against an uncontrolled out-gassing of the detector materials. The use of moving pots also allows to put detectors in a safe position when conditions of not stable beams are present. Figure 1.6 (left) shows one of the two RP stations which are installed on both sides from the interaction point IP5 on the beam pipe of the outgoing beam at a distance of about 147 m and 220 m. These positions, chosen according to the constraints given by
the space available among the LHC machine components, allow to obtain the parallel-
to-point focusing condition in both x and y directions for the RP at 220 with a beam
optics characterized by $\beta^* = 1540$ m, while for $\beta^* = 90$ m that condition is fulfilled only
in the y direction. A magnetic dipole between the two RP stations provides a magnetic
spectrometer allowing an accurate proton momentum reconstruction.

![Figure 1.6: Left: one TOTEM Roman Pot station. Right: arrangement of silicon detectors inside
two vertical and one horizontal pots at a RP unit.](image)

The silicon detectors inside the pots are characterized by “edgeless” technology: they
have been properly designed in order to reduce the insensitive edge area facing the beam to
only about $50 \mu m$ by introducing a current terminating structure [7]. This feature, together
with the possibility to move the detectors at $\sim 1$ mm from the beam axis, is relevant in
order to detect protons elastically scattered at angles down to few $\mu rad$ and, consequently,
to reduce the error on the total cross section measurement due to the extrapolation of the
elastic cross section to the optical point (see chapter 2). For the same reason the stainless
steel bottom foil of the pot (the one facing the beam) has been reduced to a thickness
of $150 \mu m$, so to occupy the minimal space between the beam and the silicon detectors.
Each RP station is composed of two units (see fig. 1.6, left), each of which consists of
three pots. Two of them approach the beam vertically from the top and the bottom and
one horizontally which allows to reconstruct the $\xi$ of diffractively scattered protons (see
fig. 1.6, right). Each pot is equipped with a stack of 10 planes of edgeless planar silicon
strip detectors. Half of them have their strips oriented at an angle of $+45^\circ$ with respect to the edge facing the beam, and the others at an angle of $-45^\circ$. Each plane has 512 strips with a pitch of 66 $\mu$m, allowing a single hit resolution of about 20 $\mu$m. Irradiation studies on these silicon detectors have shown similar aging effects as for devices using standard voltage terminating structures. It is expected that the present detectors will probably be working up to an integrated luminosity of about 1 fb$^{-1}$.

### 1.2.2 T1 telescope

The T1 telescope [1] will be installed in the CMS End Caps between the vacuum chamber and the iron of the magnet, at a distance of 7.5 to 10.5 m on both sides of IP5. Each T1 telescope arm (see fig.1.7, left), covering a pseudo-rapidity range of $3.2 < |\eta| < 4.7$, is made of five planes, equally spaced in $z$, each one consisting of six trapezoidal Cathode Strip Chambers (CSC). The Cathode Strip Chamber is a multi-wire proportional chamber with segmented cathode read-out. Even if this kind of gas detectors are slow, the response time for a CSC with a 10.0 mm gas gap (Ar/CO$_2$/CF$_4$ 40%/50%/10% is used) is still compatible with the expected hit rates for TOTEM. The detector sextants in each plane are rotated with respect to each other by angles varying from $-6^\circ$ to $+6^\circ$ in steps of $3^\circ$, in order to improve the pattern recognition for track reconstruction and to reduce the localized concentration of material in front of the CMS HF calorimeter. The printed
boards for the two sides of a CSC are identical, and give an assembled detector with cathode strips having an angle of $\pm 60^\circ$ with respect to the orientation of the anode wires which give radial coordinate measurement. Beam tests on final prototypes have shown a spatial resolution of about 0.8 mm when using a digital read-out. Aging studies, performed at the CERN Gamma Irradiation Facility, have shown no loss of performance after an irradiation resulting in a total charge integrated on the anode wires of 0.065 C/cm, which corresponds to an accumulated dose equivalent to about 5 years of running at a luminosity of $10^{30}$ cm$^{-2}$s$^{-1}$. Table 1.2 summarizes the basic parameters of T1 CSC chambers.

<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: T1 CSC chamber basic parameters [1].

1.2.3 T2 telescope

The T2 telescope [8] is located starting at $\pm 13.5$ m on both sides of IP5 and detects charged particles in the pseudo-rapidity range of $5.3 \leq |\eta| \leq 6.5$. Thanks to the good $\eta$ resolution in track reconstruction, down to 0.04 (see chapter 3), it is expected to give a good capability in discriminating against possible beam-gas background or secondary particles produced in interactions with the beam pipe. The T2 telescope is based on “Gas Electron Multiplier” (GEM) technology, which is characterized by high rate capability, good spatial resolution and good resistance to radiation. These characteristics are very important, as the T2 telescope is expected to stand up to a rate (including background) of 1 MHz cm$^{-2}$ at a luminosity of $\mathcal{L} = 10^{33}$ cm$^{-2}$s$^{-1}$.

T2 GEM chambers

GEMs are gas-filled detectors invented a decade ago by Fabio Sauli [9] and already successfully used in COMPASS and LHCb experiments. This kind of gaseous detector is
1.2 Overview on TOTEM detectors

able to detect charged particles with very high efficiency and soft photons with quite low efficiency\(^7\).

The T2 GEMs [8] use the same baseline design as the one adapted in COMPASS with a “triple-GEM” structure (fig. 1.10, left): three GEM foils are used in cascade in order to achieve a high gain (~ 8000), reducing at the same time the discharge probability below \(10^{-12}\). Each GEM foil consists of a 50 \(\mu m\) polyimide plane with 5 \(\mu m\) copper cladding on both sides on which, by using conventional photo-lithographic methods, a high density of double conical holes (with a distance of 140 \(\mu m\)) is realized. The diameter of the holes used in the T2 GEMs are 65 \(\mu m\) in the middle of the plane and 80 \(\mu m\) on the surface.

Ionizing particles interact with the gas filling the chamber (Ar/CO\(_2\) 70/30) in the drift zone (see left picture on fig 1.10), producing primary electrons. An electric field of about 2.4 KV/cm carries the electrons towards the holes of the top GEM-plane, where an electric field of about 50 KV/cm is present. Such high field generates the electron multiplication (about a factor 20 for this configuration) inside the GEM channels. An electric field of about 3.6 KV/cm in the transfer zone guides the electron cloud towards the successive GEM plane and so on until the charges are collected on the readout board.

The above description reveals a typical feature of the GEM detectors: the amplification and the readout stages are fully decoupled, allowing a more powerful chamber optimization according to the T2 telescope requirements.

The read-out board, described in next subsection, has been explicitly designed for TOTEM in order to obtain the required spatial resolution. A time resolution of about ~ 18/20 ns is obtained by using the detectors electric fields reported above which also give a GEM rate capability of about one order of magnitude bigger than the limits of 1 MHz cm\(^{-2}\), which the T2 detector must sustain. Furthermore, COMPASS triple-GEM detectors aging tests have shown that a charge up to 20 mC/mm\(^2\) can be integrated on the read-out board without aging effects. This corresponds to run TOTEM for at least 1 year at luminosities of \(10^{33}\) cm\(^{-2}\) s\(^{-1}\). It is so assumed that TOTEM T2 triple-GEM can be operated during the first 3 years of LHC running. These features makes the triple GEM

\(^7\)For instance, reminding that photons in gas detectors interact mainly for photoelectric effect, a 8 KeV photon has a mean free path in Argon of 55 mm; therefore, in 3 mm of gas active zone, only 5% of events makes a photoelectron. Moreover, increasing the photon energy the mean free path increases exponentially, reducing furthermore the detection efficiency.
technology a proper choice for the T2 telescope needs.

**T2 geometry**

Each T2 telescope arm (see fig. 1.7, right) is made by 20 triple-GEM detectors having an almost semi-circular shape (see fig. 1.8) with an inner radius matching the beam pipe and combined in two half-arms approaching each side of the vacuum pipe in order to have a full azimuthal coverage. Each T2 half-arm is made by 10 aligned detector planes mounted in a “back-to-back” configuration. Actually, each semicircular-plane covers a $\phi$ range of $192^\circ$ in order to minimize inefficiencies at the detectors edges. Hence each plane has two overlapping regions of $12^\circ$ around the $y$ axis\(^8\). The detailed relative positions of the detectors are (see fig. 1.9):

- the Z distance between two semi-planes, covering the same area in the xy plane and having the same orientation with respect to the interaction point, is 86.0 mm;
- the Z distance between one semiplane and the semiplane placed on its back side is 24.6 mm;
- the Z distance between two semi-planes forming the same circular plane is 43.0 mm.

\(^{8}\)Hereafter the $z$ axis is parallel to the beam, the $y$ axis is perpendicular to the plane of the accelerator, the $x$ axis points toward the center of the accelerator.
1.2 Overview on TOTEM detectors

Figure 1.9: Relative position of the triple-GEM detectors inside one T2 telescope.

Figure 1.10: Left: transverse view of the T2 triple GEM detector. The charges released by the ionizing particle in the 3mm drift zone are amplified by each GEM foil and then collected on the read-out board. Right: part of the read-out board showing the patterns of the two separate layers.
1.3 The beam pipe

The Z position of the first GEM plane with respect to the IP is 13828 mm. Fig. 1.10 (left) shows a transversal view of each semicircular detector composed by three GEM foils put in cascade and separated by 2 mm insulator spacers. The read-out board, explicitly designed for TOTEM, has an inner radius (distance from the z-axis) of 42.46 mm and an outer radius of 144.46 mm. It is made by two separate layers with different patterns (see fig. 1.10, right): one with $256 \times 2$ concentric circular strips, $80 \mu m$ wide and with a pitch of $400 \mu m$, which covers an azimuthal angle of $2 \times 96^\circ$ and allows track radial reconstruction; the other layer is a matrix of $24 \times 65$ pads, varying in size from $2 \times 2 \ mm^2$ to $7 \times 7 \ mm^2$ (for a constant $\Delta \eta \times \Delta \phi \sim 0.06 \times 0.05 \ rad$), which provides level-1 trigger information as well as track azimuthal angle reconstruction. Beam tests on final production detectors have shown a spatial resolution in radial coordinates of about $100 \mu m$ with digital VFAT read-out.

A conical section, with an angle of $30^\circ$, reduces the diameter from 313.8 mm (reached at
the end of the EndCap region) to 170 mm. This section covers a pseudo-rapidity region from $\eta = 4.9$ to $\eta = 5.53$. Assuming a wall thickness of 2.5 mm, a particle which cross this section close to $\eta = 5.53$ finds about 5 mm of steal in its path. From 10720 mm to 13340 mm, another cone section, pointing to IP with $\eta = 5.53$, runs under the CMS Hadronic Forward (HF) calorimeter. This section ends with a stainless steel window, perpendicular to the beam pipe, which reduces the beam pipe diameter from 210 mm to 55 mm and has a thickness of 0.1 mm. Inside the beam pipe, a perforated conical copper section 0.2 mm thick (acting as a “R.F. shield”) performs the same diameter reduction. The window and the copper cone intercept all particles with $\eta$ from 5.53 to 6.88. Close to the window it is foreseen the installation of ion pumps (see fig. 1.12, bottom), which should cover a region with $\eta < 5.3$. T2 will be installed immediately behind these pumps, starting at $z = 13828$ mm with an extent of about 40 cm. The last beam pipe section of interest is the cylindrical pipe starting from the window described above and running inside the T2 telescope (see fig. 1.12). It is 1 mm thick with an inner diameter of 55 mm and intercepts particles with $\eta$ in the range: $6.88 < \eta < 6.94$ (the last value is calculated at $z=14228$ mm). Considering
1.3 The beam pipe

The beam pipe is a critical component in high-energy physics experiments, especially in forward regions. The specification of the distances from the interaction point (IP) is crucial for understanding the geometrical layout and the potential for particles to pass through various materials.

![Figure 1.12: Top: the beam pipe Forward region with the specification of the distances from IP. The location of the TOTEM T2 telescope and of the CMS CASTOR and HF calorimeters is also shown. Bottom: location of the T2 telescope, installed between the ion pumps and the CMS CASTOR calorimeter.](image)

The steel radiation length ($X_0 = 17.35$ mm) and the nuclear interaction length ($\lambda_I \sim 17$ cm) are important parameters for estimating the amount of beam pipe material (in terms of $X_0$ and $\lambda_I$) which is crossed by a particle generated at the IP as a function of its $\eta$.

Table 1.3 summarizes the results, also showing (as reference) the spread of the polar angle.
1.3 The beam pipe

<table>
<thead>
<tr>
<th>particle $\eta$</th>
<th>$t/\lambda_I$</th>
<th>$t/X_0$</th>
<th>$\langle \theta \rangle$ (mrad)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4.9 &lt; \eta &lt; 5.53$</td>
<td>$&lt; 0.03$</td>
<td>$&lt; 0.3$</td>
<td>$\sim 0.7$</td>
</tr>
<tr>
<td>$\eta \sim 5.53$</td>
<td>$&gt; 15$</td>
<td>$&gt; 150$</td>
<td>-</td>
</tr>
<tr>
<td>$5.53 &lt; \eta &lt; 6.88$</td>
<td>$0.0006$</td>
<td>$0.006$</td>
<td>$\sim 0.32$</td>
</tr>
<tr>
<td>$6.88 &lt; \eta &lt; 6.94$</td>
<td>$2.8-3.0$</td>
<td>$27-29$</td>
<td>$7.1-7.4$</td>
</tr>
</tbody>
</table>

Table 1.3: Beam pipe effective thickness (in units of radiation length and interaction length in steel) traversed by a particle generated at the IP, as a function of its $\eta$. The multiple scattering angle (for a $\pi^-$ of $E = 10$ GeV) is also shown as reference. The effect of the thin conical copper section inside the beam pipe at $\eta$ from 5.53 to 6.88 is neglected.

distribution $\langle \theta \rangle$ for a $\pi^-$ with $E = 10$ GeV, due to multiple scattering. The calculation of $\langle \theta \rangle$ is done by using the following formula [3]:

$$
\langle \theta \rangle = \frac{13.6\text{MeV}}{\beta c p} z \sqrt{\frac{t}{X_0}} [1 + 0.038 \ln(t/X_0)]
$$

(1.10)

where $p$, $\beta c$, and $z$ are the momentum, velocity, and charge number of the incident particle and $t$ is the effective amount of traversed material. It is thus clear that particles falling in the $6.88 < \eta < 6.94$ range, crossing more than 40 cm of steel, can start nuclear or electromagnetic shower (depending on the particle kind). The same processes will happen also for $\eta$ around 5.53 where more than 2.5 m of steel can be crossed by the particle.
1.3 The beam pipe
Chapter 2

TOTEM Physics programme

In this chapter the TOTEM physics programme is presented together with an overview of the methods that will be used in order to accomplish it. Furthermore, some additional examples of interesting physics processes which in principle can be studied with the TOTEM inelastic telescopes T1 and T2 are summarized and presented in their theoretical framework.

2.1 Overview on TOTEM Physics programme

The TOTEM experiment [1] will measure the total hadronic pp cross section at LHC with an absolute uncertainty of about 1 mb. For this purpose the measurement of elastically scattered protons down to a squared four momentum transfer\(^1\) of \( -t \approx 10^{-3} \text{ GeV}^2 \) and the measurement of inelastic processes in the forward region are required (section 2.2). The differential elastic cross section will also be measured over a wide range in \(-t\) up to \(-t \approx 10 \text{ GeV}^2\), allowing to distinguish among several theoretical models (section 2.3). An interesting diffractive physics programme, including the study of single, double and central diffraction event topologies is also scheduled together with the CMS experiment (section 2.4). In addition, several joint studies on inelastic processes in the forward region are foreseen (section 2.4 and 2.5).

\(^1\)\(t\) and \(s\) are two Mandelstam variables: given a two body scattering \(1 + 2 \rightarrow 3 + 4\), \(t = -(P_3 - P_1)^2\) is the squared four momentum transfer while \(s = (P_1 + P_2)^2\) is the square of the total center of mass energy. In the relativistic limit and for small scattering angles: \(|t| \sim (p\theta)^2\), where \(p\) is the proton momentum and \(\theta\) is the scattering angle with respect to the original beam direction.
2.2 Total cross section measurement

2.2.1 Total cross section at LHC

The hadronic interaction at LHC is expected to be largely dominated by soft processes. The description of such long distance phenomena yet lacks of a first principle theory, even if many interesting semi-phenomenological models have been developed. These models are based on very fundamental physical assumptions, but their development introduces at a certain point some phenomenological aspects which are only justified in order to describe the data.

Most of the appreciated models are developed starting from the relativistic S-Matrix formalism: they provide relativistic invariance of the scattering amplitude; crossing symmetry (which relates the scattering amplitudes of the various crossed channels of the reaction); unitarity (which follows from the conservation of probability); analyticity (which postulates that the scattering amplitude should be an analytical function of kinematic variables continued to complex value) [11].

In this framework, the Regge theory was developed. It is based on the analytical continuation of the partial scattering amplitude $A_l(t)$ to complex values of the angular momentum $l$, allowing the redefinition of $A_l(t)$ as $A(l,t)$. The theory predicts that the asymptotic behaviour of the hadronic scattering in the s-channel is determined by the poles in the complex angular momentum plane of $A(l,t$) in the t-channel. This poles, which have their position depending on the value of $t$ ($l_{\text{pole}} = \alpha(t)$), are found by interpolating the resonance in the t-channel due to the exchanging of particles with angular momentum $l_{\text{pole}}$ and mass $t$. It was found that these particles lie in linear trajectories, called “Reggeons” or “Regge trajectories”. However, in order to describe the data (as the growth of the hadronic cross section with the energy), the model needs the “ad hoc” introduction of an universal phenomenological trajectory, called “soft Pomeron”, with intercept close to 1 and the quantum numbers of the vacuum [11].

Another currently used approach for the description of hadron scattering is given by the “Eikonal model”: the scattering of two hadrons at large $s$ and small $t$ is now described in the impact parameter space by introducing the profile function $\Gamma(s,b)$ ($b$ is the impact parameter of the scattering) and trying to find some reasonable parametrization of it which
must respect the unitarity constraints. Much effort is spent in order to find a “field theory based” expression for this function: however, the merely parametrization of the data with this model is very interesting. One consolidate result, which is found when analyzing the existing data with this model, is that proton seems to become bigger and blacker (in the sense that $\sigma_{el}/\sigma_{tot}$ rises toward the maximum allowed limit of 0.5) as the scattering energy grows [11], [12].

Many other phenomenological approaches could be mentioned here, all of them giving a prediction for $\sigma_{tot}$ at LHC. An important input in the validation of these models and in the understanding of hadronic interactions will be given by the TOTEM measurement of the total cross section.

Fig. 2.1 summarises the existing measurements of the total cross section from low energies up to collider and cosmic-ray energies. The dark error band shows the statistical errors to the best fit (from which $\sigma_{tot} = 111.5 \pm 1.2_{-2.1}^{+4.1}$ mb is obtained at the LHC energy), the closest dashed curves near it give the sum of statistical and systematic errors to the best fit due to the discrepancy of the two Tevatron measurements, and the highest and lowest dotted curves show the total error bands (ranging in the 90÷130 mb interval) from all models considered [13]. This large theoretical uncertainty is due to the current lack of a fully satisfactory theoretical explanation of the cross section in low momentum transfer collisions, their description relying on phenomenological models to be tuned on existing data. It is important to remember that the data are believed to be bounded by the Froissart-Martin limit which predicts the asymptotic maximum rise of the total hadronic cross section according to [11]:

$$\sigma_{tot} < \frac{\pi}{m_{Z}^{2}} (\ln s)^{2} \quad (2.1)$$

The best fit seems to predict a $\sigma_{tot}$ energy dependence like $\log^{2} s$, so this result seems to saturate (in the sense of the maximum $\sigma_{tot}$ rise with energy which is allowed by analyticity and unitarity) the Froissart-Martin limit already at $\sqrt{s} = 14$ TeV. Anyway, the big data uncertainties do not allow to close the debate about the actual functional dependence of $\sigma_{tot}$ from energy [14]. The TOTEM experiment will measure the total cross section with an absolute error of about 1 mb. This is sufficient to discriminate at least between the $\log s$ and $\log^{2} s$ extrapolation obtained from $E < 546$ GeV data reported in fig. 2.1 [15].
2.2 Total cross section measurement

Figure 2.1: Total cross section fits from the COMPETE collaboration to all available \( pp \) and \( p\bar{p} \) scattering data [13].

2.2.2 Measurement method

TOTEM will measure the total cross section by using the so called “luminosity independent method”, which relates the total cross section only to measurable quantities and to a parameter \( \rho \) which is taken from theoretical extrapolation from lower energy data. Since no reliable measurement of the machine luminosity is available, two independent equations are necessary in order to have a closed relation for the total cross section. The first equation relates the total cross section \( \sigma_{tot} \) to the luminosity \( \mathcal{L} \) and to the measured elastic and inelastic rate \( N_{el}, N_{inel} \):

\[
N_{el} + N_{inel} = \mathcal{L} \sigma_{tot}
\]  

(2.2)

The second equation is given by the “optical theorem”, based on the probability conservation (which requires that the sum of the transition probabilities from an arbitrary initial state \( |i\rangle \) to any of the possible final states \( |f\rangle \) is 1). The optical theorem relates the total cross section to the imaginary part of the elastic amplitude calculated in the forward
2.2 Total cross section measurement

\[ \sigma_{\text{tot}} = \frac{1}{s} |Im A_{el}(s, t = 0)| \]  

Let us define \( \rho = Re(A)/Im(A) \); by using the following equation:

\[ \frac{d\sigma}{d\Omega} = \frac{1}{64\pi s^2} |A(s, t)|^2 = \frac{1}{64\pi s^2} (1 + \rho^2) (Im A_{el}(s, t))^2 \]  

it is straightforward to write, inserting equation 2.3 in 2.4:

\[ \frac{dN_{el}}{d\Omega} \bigg|_{\theta=0} = L \frac{d\sigma_{el}}{d\Omega} \bigg|_{\theta=0} = \left( \frac{N_{el} + N_{inel}}{\sigma_{tot}} \right) \frac{\sigma_{tot}^2 (1 + \rho^2)}{16\pi} \]  

As already stated, the total cross section is now expressed in terms of the inelastic and elastic rate, of the elastic rate extrapolated at the optical point \( (t = 0) \) and of the \( \rho \) parameter (which is expected to be \( \sim 0.14 \) at the LHC energy). The final form of the expression for the total cross section measurement is commonly presented in the equivalent form:

\[ \sigma_{tot} = \frac{16\pi}{1 + \rho^2} \frac{dN_{el}/dt|_{t=0}}{N_{el} + N_{inel}} \]  

This relation allows TOTEM to measure the total cross section with its detectors regardless of the luminosity value. Moreover, the previous system of equations can be also solved in order to obtain the luminosity as a function of TOTEM measurable quantities:

\[ L = \frac{1 + \rho^2}{16\pi} \cdot \frac{(N_{el} + N_{inel})^2}{dN_{el}/dt|_{t=0}} \]  

therefore TOTEM will also be able to measure the LHC luminosity.

The total cross section measurement, based on the luminosity independent method, requires the proton detection by the Roman Pot (RP) stations at very small scattering angles (\( \sim \) few \( \mu \)rad), very close to the beam axis (\( \sim 1 \) mm) and the simultaneous measurement of the total inelastic rate with the T1 and T2 telescopes. In order to do so, a special accelerator optics configuration, characterized by a high \( \beta^* \) function (see section 1.1.2), is foreseen. In fact, the detection of protons elastically scattered at such small angles\(^2\) requires that in the two proton scattering \( (p_1 + p_2 \rightarrow p_3 + p_4) \) the states of proton 1 and 3 are identical; the same for the states of proton 2 and 4.
angles requires the beam angular divergence at the interaction point $\sigma_\theta$ ($\sigma_\theta \sim \sqrt{1/\beta^*}$) to be small compared to the scattering angle itself. Moreover, the proton revealed in the Roman Pot is required to be reasonably away from the beam envelope ($\sigma_{\text{env}}$), typically at least 10 $\sigma_{\text{env}}$. In order to perform the total cross section measurement at the level of $1 \div 2\%$, the TOTEM Collaboration will exploit an optical configuration at $\beta^* = 1540$ m in which a parallel-to-point-focusing condition is reached for the RP placed at 220 m from IP. As already explained in subsection 1.1.2, this condition is important because it allows to measure the scattering angle at the IP regardless the vertex position in the transverse plane. Even if running at high $\beta^*$ reduces the luminosity, the elastic scattering at small angles has a quite large cross section, hence there will be no problem, related to the low statistics, to extrapolate its value to the optical point ($t=0$). However, it is not foreseen that the LHC will run in such beam conditions in the early stage of its functioning. A run at $\beta^* = 90$ m is indeed planned in a shorter time scale which will allow TOTEM to make the total cross section measurement (even if with $\sim 5\%$ uncertainty) one of the first physics result at LHC.

The measurement of the inelastic rate, which is necessary for the total cross-section determination, will be performed with the T1 and T2 detectors which will also provide the reconstruction of the primary vertex, so that beam-beam events can be disentangled from background events (mainly from beam-gas interactions and halo muons).

The systematic error for the measurement with $\beta^* = 90$ m 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 $\beta^* = 1540$ m measurement the total inelastic rate will give the main systematic uncertainty. In this case, the uncertainty will be dominated by trigger losses in Single Diffractive events ($\sim 0.8\%$), which will occur whenever the invariant mass of the fragmented system will be quite low (below 10 GeV/$c^2$) so that the particle pseudorapidities will be beyond the T2 tracker acceptance [16]. The theoretical uncertainty related to the estimate of the $\rho$ parameter is expected to give a relative uncertainty contribution of less than 1.2\% (considering for instance the full error band on $\rho$ extrapolation as derived in ref [13]).
2.3 Elastic cross section measurement

The study of large impact parameter collisions such as elastic scattering processes is of fundamental importance for distinguishing among different models of soft proton interactions. High energy elastic nucleon scattering represents one of the collision processes in which very precise data over a large energy range are available. Fig. 2.2 shows the differential cross section of elastic \(pp\) interactions at \(\sqrt{s} = 14\) TeV, as predicted by different models tuned on lower energy data \([17]\).

Several regions in \(t\) can be identified depending on the physics of the interaction which is involved:

- \(|t| < 10^{-5}\text{ GeV}^2\): the Coulomb region, where elastic scattering is dominated by one photon exchange, described by the Rutherford formula: \(d\sigma/dt \sim 1/t^2\).

- \(2 \times 10^{-3}\text{ GeV}^2 < |t| < 0.4\text{ GeV}^2\): the hadronic region, described in a simplified way by “single-Pomeron exchange”. The differential cross section in this region is
2.3 Elastic cross section measurement

approximately exponential: \( d\sigma/dt \sim e^{-B|t|} \).

- Between the above two regions, interference between the nuclear and Coulomb scattering, complicates the extrapolation of the nuclear cross section to \( t = 0 \).

- \(|t| > 0.4 \, \text{GeV}^2\): this region exhibits the diffractive structure of the proton. As shown in fig. 2.2 the appearance of diffractive maxima and minima recalls the distribution of the light intensity which characterize the diffraction of a plane wave by a circular disk.

- \(|t| > 1.5 \div 3 \, \text{GeV}^2\): in this region there is the domain of central elastic collisions which might be described by perturbative QCD, e.g., in terms of three gluon exchange with a predicted cross section proportional to \(|t|^{-8} \) [18].

The interference region and the beginning of the hadronic region are important for the extrapolation of the differential counting rate \( dN_d/dt \) to the optical point, therefore for the \( \sigma_{\text{tot}} \) measurement. The \( t \)-dependence of the exponential slope \( B(t) = \frac{d}{dt} \ln \frac{d\sigma}{dt} \) reveals slight model-dependent deviations from the exponential shape. This theoretical uncertainty contributes to the systematic error of the total cross-section measurement. The fit is typically performed with a quadratic polynomial parametrization in the \(|t|_{\text{min}} < |t| < 0.25 \, \text{GeV}^2\) interval, where \(|t|_{\text{min}}\) depends on the acceptance for protons elastically scattered at small angles, which is related to the beam optics (see subsection 1.1.2). The expected uncertainty on the extrapolation to the optical point will be related to \(|t|_{\text{min}} \sim 0.002 (0.04) \, \text{GeV}^2\) for a beam optics characterized by \( \beta^* = 1540 (90) \, \text{m} \).

Fig. 2.2 shows that there is a model dependence of the predictions which is very pronounced at high \(|t|\). To discriminate among different models it is thus important to precisely measure the elastic scattering over the largest possible \( t \)-region. As shown in fig. 2.2, under different beam optics and running conditions, TOTEM will cover the \(|t|\)-range from \( 2 \times 10^{-3} \, \text{GeV}^2 \) to about \( 10 \, \text{GeV}^2 \) spanning the elastic cross section measurement for over 11 orders of magnitude.
2.4 Diffractive and inelastic processes

Diffractive processes are expected to give a large contribution to the total pp hadronic cross section at LHC. Indeed, including elastic processes, diffraction will contribute for about 50 mb. Fig. 2.3 shows the typical event topology for non diffractive (Minimum Bias) and diffractive processes (Single Diffraction, Double Diffraction, “Double Pomeron Exchange”, and higher order “Multi Pomeron” processes) together with the associated cross sections, as expected at the LHC. After a general introduction to diffractive processes given in subsection 2.4.1, a hard diffractive process which could be studied by involving the TOTEM inelastic telescopes, the proton diffraction into three jets, is described in subsection 2.4.2. The importance of inelastic and diffractive processes measurement for high energy cosmic ray physics is motivated in subsection 2.4.3.

2.4.1 Diffractive processes

Diffractive processes are subdivided in “soft” and “hard”. Soft diffraction gives almost the overall contribution to the diffractive cross section; hard diffraction is very interesting because of the introduction of a hard scale in the process. Essentially, the difference is in the presence of jets in most of final states of hard diffraction. The main characteristic of diffractive processes are summarized in the following points [11]:

![Event Topology Diagram](image-url)
2.4 Diffractive and inelastic processes

- A process is diffractive if there is a large rapidity gap in the final state phase space which is non-exponentially suppressed. A rapidity gap is a region of pseudo-rapidity devoid of particles and it is a typical signature for diffractive processes. Quite generally, “large” means a rapidity gap greater than 2 and “non-exponentially suppressed” means that the probability of finding the gap in the final state is not a strong function of the gap width.

- Diffractive reactions are supposed to happen without exchange of quantum numbers between the colliding hadrons.

With respect to the first point, an example is given by Single Diffractive processes at HERA \((e + p \rightarrow e + p + X)\). Calling \(M_x^2\) the invariant mass of the fragmented system \(X\), the rapidity gap \(\Delta \eta\) between the proton and the system \(X\) is related to \(M_x^2\) via the following equation\(^3\):

\[
\Delta \eta = - \ln \left( \frac{M_x^2}{s} \right)
\]

In fig. 2.4 the distribution of \(\ln (M_x^2)\) for Single Diffractive candidate events found at HERA is reported: the flat zone of the distribution of fig. 2.4, also predicted by phenomenological models, is the contribution of Single Diffractive events.

The second point listed above leads theoreticians to study diffractive processes in terms of pomeron exchange. Big efforts have been dedicated in order to understand the nature of pomeron interaction. In hard diffractive processes the pomeron is identified as colorless gluon ladder exchanged by partons; however, there is not yet a satisfying theory which explains all the aspects of this kind of hadronic processes.

The majority of diffractive events exhibits intact (“leading”) protons in the final state, characterized by their \(t\) and by their fractional momentum loss \(\xi \equiv \Delta p/p \sim \frac{M_x^2}{s}\), most of which (depending on the beam optics) can be detected in the TOTEM RP detectors. Already at an early stage, TOTEM will be able to measure \(\xi\)-, \(t\)- and mass-distributions in soft Double Pomeron and Single Diffractive events. The integration of TOTEM with the CMS detector will offer the possibility of more detailed studies of the full structure of diffractive events, with the optimal reconstruction of one or more sizeable rapidity gaps in the particle distributions which can be obtained when the detectors of CMS and

\(^3\)In the approximation of \(s, M_x^2 \gg 1 \text{ GeV}^2\)
2.4 Diffractive and inelastic processes

Figure 2.4: Distribution of $\ln(M_x^2)$ (where $M_x$ is the invariant mass of the fragmented system X) in Single Diffractive processes at HERA [19]. According to eq. 2.8, this observable is proportional to the rapidity gap between the proton and the fragmented system.

TOTEM will be combined for common data taking, as detailed in ref [20]. For this purpose the TOTEM triggers, combining information from the inelastic detectors and the silicon detectors in the RPs, are designed to be also incorporated into the general CMS trigger scheme.

2.4.2 Proton diffraction into three jets

One interesting QCD prediction for hard processes dominated by large longitudinal distances is that if a hadron is found in a small size configuration of partons, the interaction with the target has a small cross section. It results that a sufficiently energetic wave packet (projectile) with zero baryon and colour charges, localized in a small transverse volume in the impact parameter space, can be described by a $q\bar{q}$ pair [21]. In order to assume the transverse dimension of this dipole “frozen” during the target crossing, it is necessary
that the interaction takes place only with the low-\(x\) partons of the target (described by a certain distribution \(G_N\)); so the gluon contribution is expected to dominate because the gluon PDF growing at low \(x^4\) [22]. If the above conditions are satisfied, the cross section for the dipole interaction with the target goes as \(\sim d^2 x G_N(x, Q^2)\) where \(Q^2 \sim 10/d^2\) and \(d^2\) is the transverse size of the dipole [23]. This kind of small \(x\) processes are known as “colour transparency” phenomena, to remind the transparency of hadronic matter to the propagation of spatially small colour-singlet configurations.

An important example of these processes, already observed at FNAL, is the dissociation of the pion into two jet [24]: \(\pi + A \rightarrow 2\text{jet} + A\). The experimental data are explained assuming the state of the meson composed by its two valence quark localized in a small transverse area (minimal Fock space configuration). Similarly, at LHC it is expected to find protons having a significant amplitude in their \(|uud\rangle\) minimal configuration state. Because of the small interaction of the “proton dipole”, the proton-target remains intact after the collision while the projectile comes outside its mass-shell. The excited projectile has now the possibility to fragment into three quark-jets, with increasing average jet \(P_T\) as the transverse dipole area decreases. Calling \(z_1, z_2, z_3\) the longitudinal momentum fractions of the three quark inside the projectile and \(P_{T1}, P_{T2}, P_{T3}\) the transverse momenta of these quarks, the differential cross section of the process has the following form:

\[
\frac{d\sigma}{dz_1 dz_2 dz_3 dP_{T1}^2 dP_{T2}^2 dP_{T3}^2} \sim \frac{|z_1 z_2 z_3|^2}{P_{T1}^4 P_{T2}^4 P_{T3}^4}
\]

A numerical estimate for this process at LHC energy, assuming one quark \(P_T > 10\text{ GeV/c}\) and integrating over the other two transverse momenta, gives:

\[
\sigma(pp \rightarrow 3\text{jet} + p) = (1 \div 0.1) \left(\frac{10\text{GeV}}{P_T}\right)^8 \text{ nb}
\]

In general, a good fraction of these events (a quantitative estimate depends on the minimum \(P_T\) allowed for this process) is expected to fall in the acceptance of the TOTEM inelastic telescopes and hence could be reconstructed (maybe also using the energetic information from CASTOR). However, preliminary Monte Carlo studies [25] show that the

\[\text{Hereafter, the hadron momentum fraction carried by the parton is referred as “x” while “Q^2” is the typical interaction scale of the partonic process or the virtuality of the interaction-mediator.}\]
most probable configuration of the final state has one high $P_T$ jet and two close jets with lower transverse momenta balancing it in $P_T$. Therefore, for this process it will be highly probable to reconstruct only two jets in the T1 and T2 detectors. Despite the small cross section, the reconstruction of these processes has some facilitation. For instance, the final state should present jets with small area which are simpler to recognize (because of the Lorentz boost, the more energetic is the original parton, the smaller is the jet area). Moreover, Single Diffractive processes ($p + p \rightarrow p + X$), which could give a large background, should not have clusterized particles showing a definite azimuthal correlation so that a proper requirement on jet angular correlation should reject this. Other background rejection strategies, which take into account the $\xi$ of the unfragmented proton, could also be implemented. Preliminary studies suggest that the unfragmented proton has a very small $\xi$ very small ($\xi < 10^{-4}$ for $10^{-3} < |t| < 1$ GeV$^2$). This means the absence of the unfragmented proton in the RP detectors at 220 m when using nominal LHC optics ($\beta^* = 0.5$ m) or early LHC optics ($\beta^* = 2$ m). When using $\beta^* = 90$ m, it is possible to detected the unfragmented proton but a large fraction of events are still lost. The request of the absence of protons in the RP could reject a certain fraction of SD events (for which the relation $\frac{d\sigma}{d\xi}_{SD} \sim \frac{1}{\xi}$ holds). These aspects surely need a more accurate study as they could be important in order to disentangle as much background as possible from the 3 jet proton dissociation events. It is anyway clear that such kind of processes can be studied only if a proper jet finding algorithm, based on T1 and T2 tracking algorithm informations, is available. This gives the main motivation for part of the work developed in this thesis and reported in chapter 4.

2.4.3 Tuning of cosmic ray Monte Carlo generators

The accurate information on the basic properties of $pp$ collisions provided by TOTEM will also give a significant contribution to the understanding of very high energy cosmic ray physics. A challenging subject in astrophysics is represented by primary cosmic rays in the PeV ($10^{15}$ eV) energy range and above. The rate of cosmic particles above 100 PeV is about $10^{-4}$ particles per m$^2$ per year, too low for a reliable quantitative analysis. A center of mass energy of 14 TeV correspond to 100 PeV energy for a fixed target collision in the air, which will be provided by LHC with a very high event rate. Several high energy hadronic
interaction models are nowadays available describing the nuclear interaction of primary cosmic ray entering the upper atmosphere and generating air showers. They predict energy flow, particle multiplicity and other quantities of such showers which characteristics are related to the nature of the primary interaction and to the energy and composition of the incident particle. However, there are large differences among the predictions of currently available models, with significant inconsistencies in the forward ($|\eta| > 5$) region. For instance, fig. 2.5 shows the charged particle multiplicity distribution for inelastic non-

diffraction (left) and diffractive (right) events as a function of pseudorapidity, as obtained with different Monte Carlo generators. Consequently, the study of inelastic processes in pp collision at LHC performed by TOTEM T1 and T2 trackers and by CMS CASTOR calorimeter will be very important to validate/tune these generators [20]. Several quantities can be measured by TOTEM and CMS and compared with model predictions, among which: energy flow, elastic/total cross section, fraction of diffractive events and particle multiplicity.
2.5 Low $x$ physics

The opportunity of particles detection at high $\eta$ allows to investigate some important feature of QCD dynamics involving partons having small value of hadron fractional momentum (commonly referred as $x$) in the hadron wave-functions. For instance, the measurement of forward jets is important in order to gain informations on the proton parton distribution function (PDF) at low $x$. An example which shows the connection between low-$x$ and forward physics is in fact given by the forward Di-jet production. For a $2 \rightarrow 2$ parton scattering, the parton $x_1$, $x_2$ values before scattering are related to the outgoing parton rapidities $y_3$, $y_4$ by the following equation [20]:

$$
x_1 = \frac{1}{2} x_T (e^{y_3} + e^{y_4}) \quad x_2 = \frac{1}{2} x_T (e^{-y_3} + e^{-y_4})
$$

(2.11)

where $x_T = 2 k_T / \sqrt{s}$, $k_T$ being the outgoing parton transverse momentum. From this equation, it follows that Di-jet measurements in the same hemisphere of T2/CASTOR acceptance ($5.2 < |\eta| < 6.5$) would probe a $x$ value down to $10^{-6}$. This would allow to put constraints on the PDF at small $x$ by comparison between data and theoretical model. The recognition of such events could be accomplished by using a combined information from CASTOR hadronic/electromagnetic calorimeter ($5.3 < |\eta| < 6.6$)\(^5\) and the T2 tracking detector ($5.3 < |\eta| < 6.5$). It is important to remember that CASTOR has only azimuthal segmentation (16 sectors, each of about $22^\circ$), hence it is necessary to use information from T2 about the charged particle polar coordinate in order to best reconstruct this kind of forward events. Indeed, in this thesis work it is also investigated the possibility to find jets by using only the charged particles topological information as reconstructed with the T2 telescopes.

As reported in a dedicated subsection, also forward Drell Yang processes are very important for low $x$ physics study. Also for the analysis of this kind of events the combination of T2 tracking with CASTOR calorimetry will be crucial.

\(^5\)At an early stage of LHC running only one arm of CASTOR calorimeter will be installed on one side of IP5
2.5 Low x physics

2.5.1 PDF saturation

Just for completeness and because of the current interest about this topic, it is opportune to remind some basic aspects about PDF saturation. One of the most significant discoveries at HERA is the strong growth of the inclusive DIS cross section for decreasing Bjorken-$x$ at fixed photon virtuality $Q^2$ as well as for increasing $Q^2$ at fixed low $x$. In particular, comparisons of theory with data confirmed a fast growth of gluon density when $x$ decreases. However, at some small value of $x$, the distribution blow-up should be saturated otherwise there would be an unitarity violation. This is not in contradiction with the prediction of the standard evolution equations (DGLAP for the evolution with $Q^2$ and BFKL for the evolution with $x$ [22]) because in this domain they should not be applicable anymore since they can not account for gluon recombination processes due to the high gluon density. Therefore, when fixing the virtuality of the probe $Q^2$ which can be related to the “area” $r^2$ occupied by the gluon via $r^2 \sim 1/Q^2$, there should be a $x_{\text{max}}$ such that below this value the gluon density enters in the saturation regime. Similarly, fixing $x$ and decreasing $Q^2$, a value $Q_{\text{min}}$ is reached for which the gluon distribution starts to saturate. A representation of the partons inside the proton in the $\log(1/x) - \log(Q^2)$ plane is shown in fig. 2.6 [26], where the perturbative equations which govern the PDF evolution are also shown in their regions of application and the saturation region is delimited by the dotted line. Big efforts are spent in order to find models which establish the $x - Q^2$ saturation region and predict the features of the saturated regime.

In the following subsections some particular processes involving low-$x$ partons, and which could be of interest for the TOTEM inelastic detectors, are presented.

2.5.2 Minijet

Another possible measurement involving T2 regards events with minijet production (a minijet is a jet having low transverse momentum) [27]. The minijet inclusive cross section
Figure 2.6: QCD “phase diagram” in the (log(1/x), log(Q^2)) plane. Different parton evolution regimes (DGLAP, BFKL, saturation) are indicated.

\[ \sigma_{2 \rightarrow 2}^{inc}(s, p_t) = \sum_{i,j,k,l} \frac{K}{1 + \delta_{kl}} \int dx_1 dx_2 \int dp_{t1}^2 \frac{d\sigma_{i,j \rightarrow k,l}}{dp_{t1}^2} f_i(x_1, \mu^2) f_j(x_2, \mu^2) \theta(p_t - p_{t1}^c) \]  

(2.12)

where \( f_i, f_j \) are the PDF for the two ingoing partons, \( \mu^2 \) is the hard scale of the process, \( K \) is a constant factor used to correct for higher order terms.

At the LHC energy the cross section (2.12) is large and it dramatically depends on the minimum minijet transverse momentum cut \( (p_{t1}^c) \) assumed in the integration of this equation. In order to be consistent with inelastic unitarity bounds, the comparison with already existing models for the total inelastic cross section at LHC energies imposes to only use minijet \( p_t \) greater than 2.5 GeV/c and to consider only peripherical collisions with impact parameter \( b > 1.5 \) fm. With this assumptions, fig. 2.7 shows a countorn plot of the integrand of the previous equation as function of the rapidities of the two minijets.

The red solid curve shows the values where the integrand is half of its maximum.

\[ \text{Leading twist} \]

The word “leading twist” regards the approximation used for the parton distribution function evolution [28]; for 2 \( \rightarrow \) 2 parton scattering, in the leading twist approximation, all the gluon ladders joining the two partons are summed and leading order splitting function are used.
Hence the important region for minijet production also extends to the limit $|y_1| \sim 6$ and $|y_2| \sim 4$, corresponding to probed $x$ of $x \sim 4 \cdot 10^{-6}$ and $x \sim 0.08$ with at least one minijet in T2. Moreover, in the corner regions bounded by the solid blue lines, which contain both partons having high $|\eta|$ in the same hemisphere (and therefore one of the parton $x$ is very small, see 2.11), the effect of the gluon density saturation at small $x$ will be important. The blue edges delimitation is found according to the Golec-Biernat and Wusthoff (GBW) model and the solid blue line is the result of rescaling the saturation scale by a factor of $9/4$ necessary for $gg \to gg$ scattering.

The importance of jet reconstruction in the very forward region is again evident.

### 2.5.3 Forward Drell-Yan

Another example of interaction which probes very small $x$ values and which could be of interest for the T2 detector is the forward Drell-Yan process. The Drell-Yan process is the
2.5 Low x physics

electroweak production of a lepton pair from the annihilation of a quark-antiquark pair. If a quark-antiquark pair with fractional momenta $x_1-x_2$ annihilates in a $pp$ collision, the virtual photon/$Z_0$ (or the dilepton system) rapidity $Y$ and invariant mass $M$ are related to $x_1-x_2$ via the following formulae [29]:

$$M^2 = s x_1 x_2 \quad \text{and} \quad x_{1,2} = \frac{M}{\sqrt{s}} e^{\pm Y}$$

(2.13)

with $\sqrt{s} = 14$ TeV, the centre-of-mass energy of the colliding protons. Boosting the lepton pair to large rapidities allows to be sensitive to the low $x$ regime of the sea (anti)quark distributions.

As CASTOR provides a limited azimuthal angular information, while not providing polar angular information, the T2 tracker is necessary in order to measure the dilepton mass $M$ and the fractional momenta $x_1-x_2$, in events triggered on electromagnetic energy deposits detected by CASTOR.

Figure 2.8: Distributions of $x_{1,2}$ (top) and energy (bottom) for Drell-Yan electrons with invariant mass $M_{ee} > 4$ GeV/c$^2$, both within the acceptance of T2/CASTOR [20].

The effect of saturation should manifest as a decrease of about 30% on the Drell Yang cross section in the T2/CASTOR acceptance. The cross section expected for the
production of a Drell Yang $e^+e^-$ pair with $M_{ee} > 2 \text{ GeV}/c^2$ is estimated to be 2.6 nb (if the PDFs saturation is not taken into account) when the lepton pair is produced within the acceptance of T2/CASTOR. Fig. 2.8 (top) shows the kinematic coverage as a function of $\log(x)$ for Drell-Yan events with invariant mass $M_{ee} > 4 \text{ GeV}/c^2$ and with both electrons within the T2/CASTOR acceptance [20]. Values of the order of $x \sim 10^{-6}$ are reached at large rapidities. The corresponding distribution for the electron energies is also shown in the bottom plot. The $e^+e^-$ pair in T2/CASTOR tends to have large energies and can be selected by requiring $E_e > 300 \text{ GeV}$.

### 2.5.4 Mueller Navelet jets

One of the important questions still open in QCD is how to describe the scattering amplitudes in the Regge limit, where the interacting parton center of mass energy is much larger than all other Mandelstam invariants and mass scales [30]. In this region, divergent logarithms at small-$x$ are resummed using the Balistky-Fadin-Kuraev-Lipatov (BFKL) evolution equation. An example of processes where BFKL effects should be dominant is the “Mueller Navelet jets” events, where two jets with large and similar transverse momenta $k_1, k_2$ and quite large fractional momentum $x_1, x_2$ are separated by a big rapidity

![Figure 2.9: Azimuthal correlations between jets with $\Delta \eta = 6, 8, 10$ and $11$ and $p_T > 5 \text{ GeV}/c$. This measurement will represent a clear test of the BFKL regime [31].](image)
interval $\Delta Y \sim \ln(x_1 x_2 s/k_1 k_2)$ [32]. The large rapidity separation enhances the available phase space for BFKL radiation making the Mueller Navalet process ideal in order to test the BFKL regime. A typical observable to look for BFKL effects is the measurement of the azimuthal correlations between both Mueller Navalet jets [31]. The DGLAP prediction is that the distribution of the jet azimuthal separation should peak around $\pi$, i.e. the jets should be back to back, whereas multi-gluon emission via the BFKL mechanism should lead to a smoother distribution. Fig. 2.9 shows the azimuthal separation $\Delta \phi = \pi + \phi_1 - \phi_2$ of the two jets, according to the BFKL dynamics, obtained with $k_1 = k_2 > 5$ GeV/c. As shown in this figure, the decorrelation increases with the jets rapidity separation.

Mueller Navalet jets can be a very nice measurement to be performed with the TOTEM T2 tracker and the CMS CASTOR calorimeter, as a rapidity separation $\Delta Y \sim 12$ between the two jets in the opposite detector arms can be detected. This process represent another examples showing the importance of forward jet reconstruction.
Chapter 3

Track reconstruction with the T2 detector

In this chapter the algorithms developed for track reconstruction in the T2 detector and their performances are presented. An overall description of the simulation and reconstruction chain in the CMS framework is first reported, followed by a detailed description of the software which has been developed in this thesis. Being this reconstruction software still in validation, future improvements which will allow to achieve better reconstruction performances are expected.

3.1 Simulation and reconstruction chain

From the point of view of the software structure, in order to allow a full compatibility of TOTEM and CMS data processing in future analyses studies, physics and simulation events will be reconstructed and recorded in TOTEM and CMS in the same software framework. The CMS experiment has developed a C++ based framework (CMSSW) [33] which implements a software bus model where there is one executable, called cmsRun, and many plug-in modules which run algorithms. The software for the TOTEM experiment is being developed within this framework. The CMSSW executable is configured at run time by user’s job-specific configuration files. These files tell cmsRun which data to use, which modules to run, which parameter settings to use for each module and in what order
to run the modules.

All informations about the real data acquisition or from physical events simulation are organized according to an Event Data Model (EDM). In software terms, an 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. As the event data is processed, products (generated by producer modules) are stored in the Event as reconstructed data objects. During processing, data are passed from one module to the next one via the Event and are only accessed through the Event. The Event thus holds all data taken during a triggered physics event (or generated by a MC) as well as all the information elaborated from the taken/generated data. All objects in the Event may be individually or collectively stored in ROOT format files, and are thus directly readable in ROOT [34].

The event simulation and reconstruction are summarized in the following steps:

- **Event generation**: in the present work, this step is handled either with PYTHIA parton shower MC generator [35] or with a Particle Gun. PYTHIA allows to can generate the final state products of a 14 TeV pp collision for a wide variety of physical processes (from non diffractive inelastic (Minimum Bias) to diffractive processes; from Di-jet to Higgs mechanism at LO in PQCD)\(^1\). The Particle Gun allows to generate single or multiple particles with flat \((\eta, \phi, \text{energy})\) distributions as well as with fixed \((\eta, \phi, \text{energy})\) values at the IP.

- **Evolution of particles from IP**: the simulation of the interactions of the generated particles travelling from IP toward the detectors is handled by GEANT4 [36], which simulates the electromagnetic and hadronic interaction of the particles with matter as well as the effect of the magnetic field. In order to properly simulate the interaction of the particles with detector materials, all the detector components are described in detector geometry files. The geometry is written in Detector Description Language (DDL) in XML format files. GEANT4 will trace the interaction of the particles with the sensitive volumes described in these geometry files.

- **Digitization**: the electrical response of a given detector (for instance of a T2 GEM

\(^1\)It’s worth to remind there are other MC generator available of which Herwig is a commonly used example.
chamber), caused by the crossing of a charged particle, is simulated in the digitization step. Given the GEANT4 simulation of particle entry/exit points in/out the detector active volume and their energy deposition, proper CMSSW modules reproduce the electrical output signal of the related detector. For what concerns the T2 triple GEM detectors a module has already been developed which reproduces the digital output signal of the chambers [37]. The outputs of the digitization are the pad/strip digital status for each telescope plane. This digitization has been derived from a detailed simulation of the GEM chambers [38]: dedicated software tools (not included in the CMS framework) allow to simulate the gas ionization and electron diffusion in the drift zone, the electric field in the GEM plane, the electron multiplication in GEM holes, the diffusion in the transfer zones and the signal induction on the read-out plane (as reference see fig. 1.10, left). From studies on this detailed simulation, it was found a simple model, implemented inside the CMSSW framework, which describes the electron cloud and the charge fraction collected in pads and strips as function of the position of the primary ionization.

- **Clusterization**: a particle traversing a detector device typically fires more than one read-out channels. Consequently, proper clusterization algorithms need to be developed in order to reconstruct at best the particle position in the detector. For what concerns the clusterization of the T2 GEM strip/pads signal a CMSSW producer module has been developed in this thesis starting from a preliminary clusterization algorithm already implemented for the test beam data acquisition and analysis [39]. This module allows to properly load the digitization information necessary to compute the cluster: one pad/strip cluster is composed respectively of neighbouring pad/strip having their digital status ON. The output of this module is a collection of pad and strip clusters. The cluster informations, which are provided by this module and that will be used later, are: the cluster type (pad or strip type cluster); the number of pads/strips in the cluster; the detector ID to which the cluster belongs; the radial and angular center of gravity of the cluster (in mm and degree respectively, according to CMS convention for coordinates\(^2\)); the angular and radial uncertainty of

---

\(^2\)The coordinate system adopted by CMS has the origin centred at the nominal collision point inside the experiment, the y-axis pointing vertically upward and the x-axis pointing inward toward the center of...
the particle position associated to the considered cluster. It is important to remark that hereafter (also including the clusterization module) the developed software will be able to run on both simulated and real data.

- **Hit Reconstruction**: the cluster informations have then to be combined with the related detector position in order to give a 3D reconstruction of the particle “hit” in that detector. For the reconstruction of particle hits in T2, a proper CMSSW module has been developed in this thesis work. This module reads from the Event the collection of the clusters previously found in each plane. It aims to combine the radial and angular informations from the overlapping clusters (or single cluster if no overlap occurs) in order to have the particle position information with the smaller uncertainty possible. A ionizing particle crossing the detector planes typically switches on a group of neighbouring strips (which form a strip cluster) and one or two pads (which form a pad cluster) most of the times (see fig. 3.1). These two clusters overlap in the detector read-out plane, hence our best information about the real position of the ionizing particle in the transverse plane is found by using the angular information given by the pad cluster and the radial information given by the strip cluster. Adding the Z position of the related detector plane and the errors on the r, $\phi$ and Z coordinates, the basic informations on the particle position are thus obtained. This set of informations is called “reconstructed hit” and the output of the hit reconstruction module is the collection of all the reconstructed hits in the event. This package will be described into details in section 3.2.

- **Road finder**: The aim of this module, developed for the work of this thesis, is to group the hits found in the previous step in one or more hit collections to be successively used for track reconstruction. This step must be accomplished in order to group to the best all the reconstructed algorithm hits which, being related to the same ionizing particle, belong to different detector plane. The groups obtained

\[ \text{3.1 Simulation and reconstruction chain} \]

---

3 The distribution in number of pads and strips switched on by a traversing particle in the event simulation are still to be tuned in the digitization in order to proper reproduce the test-beam results and a proper simulation of the noise is not yet implemented.
3.1 Simulation and reconstruction chain

Figure 3.1: Strip (left) and pad (right) cluster-size as obtained from test-beam data (top) and simulation result (bottom). The simulation still needs to be tuned in order to properly reproduce the test-beam results.

In this way are called “roads”. With the exclusion of rare cases (for example, delta rays production in the ionizing GEM gas), a set of quasi-collinear reconstructed hits is expected to be found in a road associated to each ionizing particle. For each event, the main output of this module is a collection of roads, that is a collection of reconstructed hit vectors. More details on this module will be given in section 3.3.

- **Track reconstruction**: this module developed in this thesis reads from the *Event*
the collection of all the roads previously found. Since the T2 telescope is far away from the region where the CMS magnetic field has considerable intensity, the hits of each road are fitted with a straight line by using a Linear Least Square Methods. Furthermore, as the T2 detector represents a low amount of material the charged particle, multiple scattering in the detector volume is not taken into account. The algorithm implemented in this module is very similar to the tracking algorithm developed for track reconstruction in T1 [40]. Also the data format of the “T1-Track” and the “T2-Track” object are the same. In this way a full compatibility between the track reconstruction of the two TOTEM inelastic detectors is ensured, a feature considered useful for the future development of a common vertex finder. This module, which implementation will be detailed in section 3.5, also records all the reconstructed tracks (one for each road) in the Event. In particular, among the track parameters described in section 3.5, the track $\chi^2$ and the distance from the IP of the Z coordinates of the point of the track closer to the beam Z-axis ($Z_0$) will be important, in addition to the basics geometrical track information $\eta$, $\phi$ and their errors. The $\chi^2$ and $Z_0$ parameter will be extensively used in order to eliminate from the event tracks which are generated in secondary interactions.

3.2 Hit reconstruction

As already introduced in the previous section, a charged particle traversing a triple-GEM detector plane will generate a signal on both layers of the read-out board. Fig. 3.1 shows the typical strip and pad cluster size, as obtained from test beam data (top) and from digitization in simulated events (bottom). Signal fluctuation below the digital VFAT chip thresholds can in principle gives signals in only one layer of the read-out board: at the same time, few signals generated by electronic noise are expected to involve only one read-out layer.

In order to reconstruct the hits, the position of all the pad clusters are compared to those of all the strip clusters inside the same detector plane. The radial and azimuthal “spread” of a cluster is defined as the size of the cluster in the radial and azimuthal angle
3.2 Hit reconstruction

Calling \((\phi_{ci}, R_{ci})\) the center of the i-th cluster and \((\Delta \phi_{ci}, \Delta R_{ci})\) its spread, two clusters will form a hit of “Class 1” if such pair, belonging to the same detector plane satisfies the following overlapping condition:

\[
|\phi_{c1} - \phi_{c2}| < \max(\Delta \phi_{c1}, \Delta \phi_{c2})/2 \quad (3.1)
\]

and

\[
|R_{c1} - R_{c2}| < \max(\Delta R_{c1}, \Delta R_{c2})/2
\]

otherwise, if such overlap does not exist, the hit is obtained from only one cluster type and it is considered of “Class 2”. In what follows, only hits of Class 1 are considered; indeed, with the current model of digitization, it is found that the number of reconstructed Class 2 hits is really negligible (see fig. 3.2). However, a more accurate study about this has to be done when digitization model will be properly tuned to the test-beam data (see fig. 3.1).

The values of the hit \(R/\Phi\) coordinate (and their errors) are the ones of the strip/pad cluster related to the hit having the smaller \(\Delta R/\Delta \Phi\). It is important to remind that for pad clusters \(\delta R = \Delta R/2\) and \(\delta \Phi = \Delta \Phi/2\) while for strip clusters the radial error is defined according to the standard formula: \(\delta R = 0.4 \text{ mm}/\sqrt{12} \sim 115 \mu\text{m}\) (where 0.4 mm is the strip pitch\(^5\)) and the angular uncertainty is \(\sim 96^\circ/2\).

From the \(R-\Phi\) center of the hit, the X and Y center coordinates are derived and their errors are obtained from simple error propagation formula. Moreover, the information of the detector plane ID is also recorded inside the cluster collection in the \textit{Event}. In particular, the following informations are available:

- the detector side respects to the interaction point (1 if it is placed at \(Z > 0\), 0 otherwise); hereafter, this parameter will be called “arm”;
- the detector “plane number” (from 0 to 4)
- the detector “plane side” (1 if the detector has the drift region facing the IP, 0 otherwise)

\(^4\)For a strip cluster size = 1 the radial spread is set to 0.4 mm, the strip pitch.
\(^5\)Direct measurements performed on test beam data have shown a digital resolution in radial coordinate of about 100 \(\mu\text{m}\) with digital VFAT read-out [10]
3.2 Hit reconstruction

Figure 3.2: Reconstructed hit $\phi$ (left) and $R$ (right) as obtained firing 1000 $\mu^-$ with $E = 10$ GeV, $\eta = 6$ and $\phi = 210^\circ$. Magnetic field and beam pipe are not included in the simulation. Top: Class 1 and 2 reconstructed hits. Bottom: Class 1 reconstructed hits.

- the “half telescope” side where the detector is installed (0 if the half arm is inside the accelerator ring, 1 otherwise).

For each detector, there is an unique correspondence between the set of these four numbers and the $Z$ position of the detector itself. In fact, firing a charged particle toward a given detector, it is also possible to retrieve from GEANT4 the $Z$ position of the mid point of the GEM chamber drift region crossed by the particle (this position will be called “GEANT4 local point”). Therefore, a simple formula has been derived which, given the detector coordinates (in terms of arm, plane number, plane side, half telescope), returns its $Z$ coordinate as given by the GEANT4 local point, which is now taken as the detector
3.3 Road finder

Z position. The explicit formula is given by the following equation:

\[
Z = \text{arm} \times [13828.3 + (\text{plane number}) \times 86 + (\text{plane side}) \times 24.6 \\
+ (1 - \text{half telescope}) \times 43.0] \text{ mm (3.2)}
\]

where the numerical values (in mm) describe the distances between the GEM detectors and the distance between the IP and the first GEM plane (as already detailed in chapter 1). The hit DZ is set to 1.5 mm, half of the GEM drift zone width. The number of strips and pads forming the hit is also associated to it as additive information. The following table resumes all the information associated to the reconstructed hit:

<table>
<thead>
<tr>
<th>Hit geometrical information</th>
<th>R, Φ, X, Y, Z, δR, δΦ, δX, δY, δZ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hit cluster composition</td>
<td>pad number, strip number, hit Class</td>
</tr>
</tbody>
</table>

Table 3.1: Reconstructed hit informations.

Fig. 3.3 shows the distribution of the difference between the \(r\) and \(ϕ\) coordinates of the reconstructed hits and the ones of the generated particles, as obtained by firing 1000 π\(^-\) with \(E = 50\) GeV, a flat azimuthal angular distribution and \(5.3 < \eta < 6.3\). The width of the distributions is related to the radial and angular resolution in hit reconstruction. Considering the RMS a \(r\) resolution \(σ_R < 0.1\) mm and a \(ϕ\) resolution \(σ_ϕ \sim 1°\) are estimated. Since the present study is only focused on checking the performances of the clustering and hit reconstruction algorithms, the magnetic field and the beam pipe are not included in the simulation. Their effect will be systematically studied when considering the full track reconstruction.

3.3 Road finder

The ultimate goal of the reconstruction software developed in this part of the thesis is to reconstruct the trajectories of the charged particles crossing the telescope. To do so an intermediate step is needed in order to group all the reconstructed hits associated to the same particle trajectory. Such group of hits is called “road ” and this section describes the method used to find the roads in the T2 telescope. The CMSSW module developed for this
Figure 3.3: Left: difference between the reconstructed hit radial position in a given detector and the expected radial position of the particle in the same detector active zone. Right: difference between the reconstructed hit $\phi$ and the generated particle $\phi$.

Purpose is based on a “histogramming method” [41] which clusterizes the reconstructed hits in the $R - \phi$ plane. This method can be summarized as follows:

1. The $R - \phi$ plane corresponding to the telescope area ($R_{\text{min}} = 42.46$ mm, $R_{\text{max}} = 144.46$ mm, $\phi_{\text{min}} = 0^\circ$, $\phi_{\text{max}} = 360^\circ$) is divided in a grid of rectangular cells of opportune size.

2. Each reconstructed hit is assigned to the cell containing the ($R$-$\phi$) hit coordinates. When all hits are assigned, the number of hits falling inside each cell is computed.

3. A minimum value of the cell multiplicity is assigned and the list of all cells with multiplicity above that threshold (called “seeds list”) is found.

4. The list is ordered according to cell multiplicity and a second list of seeds is obtained from the first one by excluding seeds placed at less than 1 $\Phi$ and 1 R grid size from a seed with a higher multiplicity. Each road is then obtained as the collection of hits falling inside a 3x3 rectangular group of cells centred on each seed which remains in...
the seed list. The choice of the grid size, of the minimum cell multiplicity and of the grouping criteria around the seed is surely a delicate issue. Future studies could be necessary in order to optimize the road identification, in particular in a high track density topology. To facilitate this job, the road finder module has been implemented with the possibility to set at runtime the grid size and the minimum cell multiplicity, via a configuration card editable by the user. The grid used for the present study has cell size of 2 mm and 3° respectively in the r and φ coordinate. We motivate this choice looking at figures 3.4 and 3.5.

These histograms are obtained without the simulation of beam pipe and of the magnetic field in order to have an unbiased information, independent on the particular event chosen. In particular, fig. 3.4 shows in the y-axis the distribution of radial coordinates of the reconstructed hits for primary particles (µ) coming from the vertex in the T2 acceptance at two different η. The histogram on the left shows a spread of about 5 mm, while the one in the right shows that this spread is reduced when moving at higher η from the lower acceptance η value (η = 5.4). This motivates the choice of a R-side for the cell of 2 mm (and thus a road radial width of 6 mm). The fig. 3.5 shows the hit reconstruction for a Di-jet event having both partons generated in the T2 acceptance and in the same detector arm. It is clear that it is possible to appreciate a particle separation at least of about 10°, thus a road azimuthal width of 9° is chosen corresponding to a φ-side for the cell of 3°. We can also see how the hits are characterized by the same φ value, which is expected due to the absence of the magnetic field. Proper studies have indeed shown that this happens also when the magnetic is switched on, as a consequence of the fact that T2 detector is in a region where it is negligible. The φ segmentation of the grid is so expected to be valid also when the magnetic field is taken into account.

Moreover a minimum seed multiplicity of 3 hits is required. This could be in the worst case, the number of hits in each cell for a road covering three cells in radial coordinates with a multiplicity of 9-10 hits (as explained in chapter 1.2.3, T2 telescope is composed of 10 GEM-detector planes hence an ideal road has at most 20 or 10 reconstructed hits if the particle azimuthal angle is in the overlapping detector zones or not).

\[^{6}\]The overlap azimuthal intervals are 84° < φ < 96° and 264° < φ < 276°, as explained in the chapter 1.2.
3.3 Road finder

Figure 3.4: \((R, \phi)\) coordinates of reconstructed hits as obtained by firing 1000 muons with \(E = 100\) GeV and a flat azimuthal angular distribution. Left: distribution for muons generated at \(\eta = 5.4\), right: muons generated at \(\eta = 6.0\).

Figure 3.5: \((R, \phi)\) coordinates of reconstructed hits as obtained by generating one Di-jet event with both outgoing parton in the T2 acceptance, on the same side respect to the interaction point.
3.4 Track reconstruction

As already stated, due to the absence of the magnetic field in the T2 region, it is reasonable to perform track reconstruction by fitting with a straight line. The hits in each road are projected in the XZ and YZ planes, obtaining two 2-dimensional roads on which a standard minimum linear least square method \[42\] is applied. The standard formulae which have been used are:

\[
\alpha_X = \frac{\sum x z \sum \frac{1}{(\delta x)^2} - \sum z \sum \frac{x}{(\delta x)^2}}{\sum \frac{z^2}{(\delta x)^2} \sum \frac{1}{(\delta x)^2} - \sum z \sum \frac{z}{(\delta x)^2}} \tag{3.3}
\]

\[
\sigma_{\alpha_X} = \sqrt{\frac{\sum \frac{1}{(\delta x)^2}}{\sum \frac{z^2}{(\delta x)^2} \sum \frac{1}{(\delta x)^2} - \sum z \sum \frac{z}{(\delta x)^2}}} \tag{3.4}
\]

\[
\beta_X = \frac{\sum \frac{x}{(\delta x)^2} \sum \frac{z^2}{(\delta x)^2} - \sum z \sum \frac{x z}{(\delta x)^2}}{\sum \frac{z^2}{(\delta x)^2} \sum \frac{1}{(\delta x)^2} - \sum z \sum \frac{z}{(\delta x)^2}} \tag{3.5}
\]

\[
\sigma_{\beta_X} = \sqrt{\frac{\sum \frac{z^2}{(\delta x)^2}}{\sum \frac{z^2}{(\delta x)^2} \sum \frac{1}{(\delta x)^2} - \sum z \sum \frac{z}{(\delta x)^2}}} \tag{3.6}
\]

where \(\alpha_X\) and \(\beta_X\) are the angular coefficient and the intercept of the 2-dimensional line in the XZ plane and \(\sigma_{\alpha_X}\) and \(\sigma_{\beta_X}\) are their standard errors; the sums are over the road hits. Similar formulae are used in order to make the fit in the YZ plane (just substituting \(x \rightarrow y\) in the previous equations).

Once the values of these four parameters are found, it is possible to reconstruct the following geometrical parameters of the track:

- \(R_0\): is the minimum approach distance between the reconstructed 3D track and the z axis. In general, the track and the z axis are skew with respect to each other,
however their reciprocal distance has the following expression:

\[ R_0 = \sqrt{\left(\alpha_X Z_0 + \beta_X\right)^2 + \left(\alpha_Y Z_0 + \beta_Y\right)^2} \]  

(3.7)

- \( Z_0 \): is the point along the Z axis in correspondence to the minimum approach distance of the track; the calculation leads to the following expression:

\[ Z_0 = -\frac{\alpha_X \beta_X + \alpha_Y \beta_Y}{\alpha_X^2 + \alpha_Y^2} \]  

(3.8)

- \( \theta \): is the angle of the track associated versor with respect to the z axis. As the track versor is:

\[ \hat{V} = \frac{(\alpha_X, \alpha_Y, 1)}{\sqrt{\alpha_X^2 + \alpha_Y^2 + 1}} \]  

(3.9)

it follows that:

\[ \cos \theta = \frac{1}{\sqrt{\alpha_X^2 + \alpha_Y^2 + 1}} \]  

(3.10)

- \( \phi \): is the track azimuthal angle, calculated as:

\[ \phi = \arctan\left(\frac{\alpha_Y}{\alpha_X}\right) \]  

(3.11)

In addition, the track fit \( \chi^2 \) is calculated for both the projections (\( \chi_X^2 \) and \( \chi_Y^2 \)) and the track reduced \( \chi^2 \) (\( \chi_{\text{red}}^2 \)) is obtained from the following equation

\[ \chi_{\text{red}}^2 = \left(\frac{1}{2}\right) \left(\frac{\chi_X^2 + \chi_Y^2}{N - 2}\right) \]  

(3.12)

where \( N \) is the number of hits in the road.

When performing proper studies in simulated events in order to validate the track reconstruction package, it has been found that some problems arise whenever the particle crosses the overlapping region between the two half telescopes (close to 90° and 270°). Here the hits simultaneously reconstructed in the two half arms can have different reconstructed \( \phi \) (see fig 3.6), a consequence of the limited \( \phi \) resolution in the pad read-out. Moreover, from error propagation, the \( \delta X \) uncertainties on the reconstructed hit X coordinate have a
### 3.4 Track reconstruction

Figure 3.6: Distribution of reconstructed hit in the half telescope internal (left) and external (right) to the accelerator ring for a sample of $\mu$ fired at $\phi = 270^\circ$ and $\eta = 6$.

maximum at $90^\circ$ and $270^\circ$. The track projection on the XZ plane is expected to be fitted quite badly and this decrease in quality should also have a bad effect on the $\theta$ calculation from formula 3.10. As the T2 detector is designed in order to reach a good $\eta$ resolution, it would be better to not compromise his performances because of such limits on the hit reconstruction algorithm. The solution adopted at the moment use a different fitting procedure each time the road is contained in the azimuthal interval $88^\circ < \phi < 92^\circ$ and $268^\circ < \phi < 272^\circ$. In this cases, the track $\theta$, is derived from a simple fitting in the R-Z plane using the following relation:

$$ R = a_{rz}Z + b_{rz} $$

(3.13)

where $a_{rz} = \tan \theta$ and $b_{rz}$ are the unknown parameters, while the $\phi$ coordinate is obtained as an average of the hits $\phi$. In this complementary fitting procedure it is assumed $R_0 = 0$ while $Z_0$, which is an important track parameter for the studies presented in this work, is
3.5 Tracking performance

The T2 detector performance for track reconstruction has been extensively studied in several samples of simulated events. The results are reported in this section. As expected, the effect of the beam pipe on track reconstruction is not negligible because of the production of secondary particles. The studies presented here are focalized in order to understand how well it will be possible to select primary particles tracks on real data. A reasonable set of cuts on track parameters is introduced in order to select only tracks related to primary particles. The effects of these track selection criteria are reported and preliminary results on track reconstruction efficiency and resolution are presented. Preliminary studies on track reconstruction related to secondary particles generated by neutral pions are also reported.

3.5.1 Beam pipe effect

The effect of particle interaction in the beam pipe has been checked by firing 50 GeV \( \pi^− \)'s at several different pseudorapidities in the T2 acceptance region \( (5.3 \leq |\eta| \leq 6.5) \).

As first step, fig. 3.7 (top) shows the track reconstruction pseudo-efficiency\(^8\) in the ideal situation where the beam pipe and the magnetic field are not present, in order to check that the reconstruction chain works properly. More precisely, each point of this figure has

\[
Z_0 = -\frac{b_{rz}}{\tan(\theta)}
\]

For future software developments it could be important to make more systematic studies on the overlap region, maybe extending this additional fitting procedure on a wider interval in \( \phi \) inside it. However, because this problem affects only particles having azimuthal angles in a narrow \( \phi \) sub-region inside the overlapping area, dramatic changes in the final performances are not expected.

\(^5\)Most of hadrons produced in inelastic events are charged pions (hadronically interacting with matter) and neutral pions decaying to two photons soon after their generation (which will interact electromagnetically with matter). The energy distribution of such hadrons, when falling on the T2 acceptance, is typically peaked around 50 GeV (see subsection 3.5.2).

\(^8\)The term pseudo-efficiency is hereafter considered because also secondary tracks are included in its definition. So that this quantity can in principle be bigger than 1.
been obtained by generating 1000 $\pi^-$s with a flat distribution in $\phi$ ($0 \leq \phi \leq 2\pi$) at a fixed $\eta$ (shown in the x-axis) and then evaluating the total number of reconstructed tracks which, divided by 1000, gives our definition of reconstruction pseudo-efficiency (shown in the y-axis). It is clear that the simulated T2 detector performance as well as the one of the tracking algorithm allow an intrinsic track reconstruction pseudo-efficiency very close to 100%. When including the simulation of the beam pipe it is expected, as explained in section 1.3, that charged pions can make hadronic interactions when crossing it, with a consequent rise in track multiplicity due to the additional reconstruction of secondaries. The production of secondary particles is expected to be particularly important around $\eta = 5.53$ and $\eta = 6.9$ where the amount of steel traversed by the pion can be up to $\sim 15 \lambda_I$ and $\sim 3 \lambda_I$, respectively. This effect is confirmed in fig. 3.7 (bottom) where the track reconstruction pseudo-efficiency is plotted as a function of the generated particle $\eta$ when also the beam pipe and the CMS magnetic field are included in the overall detector simulation. Due to the large value of interaction lengths crossed by the particles around $\eta = 5.53$ and $\eta = 6.9$, the two spikes are composed almost exclusively by tracks associated to secondary pions. In these cases the reconstructed track $\eta$ is very different from the generated $\eta$ value and the quality of the track reconstruction, given by the track $\chi^2$, is quite reduced. As shown in the right plot of fig. 3.8, obtained by firing 50 GeV pions at $\eta = 5.59^9$ (where secondary particles production is expected), the reconstructed track pseudorapidities are for a large amount outside the T2 telescope acceptance, which is possible only for particles not coming from the IP. Moreover, in such situation, it is possible to have more than one secondary track contributing hits in the same road (see section 3.3). When performing a straight line fitting by using hits belonging to different particles the goodness of the fit decreases, thus a worsening of the reconstructed track $\chi^2$ is expected. For comparison, fig. 3.8 (left) also shows the reconstructed track $\eta$ when firing pions at $\eta = 5.9$ (where the particles which come from the IP do not find relevant amount of steel in their path). Track reconstruction is clearly not affected by secondary interactions. This is confirmed in fig. 3.9 which shows the reduced $\chi^2$ $^ {10}$ distribution for $^9$The full event simulation also includes the expected smearing on the z position for the primary vertex. Consequently, particles produced at a given $\eta$ value will have a correspondent detector $\eta$ value which has a smeared distribution around the original one. This allows a pion generated very close to $\eta = 5.53$ to have some probability to be not subject to a hadron interaction in the beam pipe.
$^{10}$Hereafter, the reduced $\chi^2$ is called directly $\chi^2$. 
Figure 3.7: Mean number of reconstructed tracks for a 50 GeV $\pi^-$ generated at the pseudo-rapidity shown in the x-axis and with a flat azimuthal distribution without (top) and with (bottom) the simulation of the beam pipe and of the CMS magnetic field. The two spikes at $\eta \sim 5.53$ and $\eta \sim 6.9$ in the bottom plot confirm the production of secondary particles in the beam pipe.

all the reconstructed tracks, firing $\pi^-$ at $\eta = 5.9$ (left) and at $\eta = 5.59$ (right). In the latter case, the higher $\chi^2$ value indicates that a given road may include hits from different particles or it may not contain all of the hits related to the track (when the secondary particle $\eta$ is too small, because the track algorithm is optimized in order to reconstruct
3.5 Tracking performance

![Histograms showing track reconstruction at different η values](image)

Figure 3.8: Reconstructed track η firing 1000 50 GeV π⁻ s at η = 5.9 (left) and at η = 5.59, where the π⁻ crosses a large amount of beam pipe steel (right).

tracks from the IP).

It is important to remind that the beam pipe will also produce secondary particles from primary π⁰ which actually don’t arrive at the telescope as they decay via π⁰ → γγ soon after their production\(^\text{11}\). The photons produced from π⁰ decay can convert to a e⁺ e⁻ pair when traversing some material or even can start an electromagnetic shower if they cross a large amount of beam pipe steel. Detailed studies about secondary track generated by neutral pions will be presented in chapter 3.5.4.

### 3.5.2 Primary track selection

In order to properly reconstruct physics events, it is important to disentangle primary tracks from secondary ones. For this purpose, some strategies which allow to select only tracks coming from the IP and to reject secondary tracks, are studied in this section. The possibility of primary tracks selection turns out to be of great importance for event vertex reconstruction as well as for jet reconstruction with topological jet algorithms (see chapter

\(^{11}\)The π⁰ lifetime at rest is about $8.4 \times 10^{-17}$ s \([3]\), which allows for a 50 GeV primary π⁰ to travel a distance of about 10 μm in the laboratory frame.
4). In fact, besides the obvious biases in primary vertex reconstruction, if this selection would not be introduced, one particle producing almost randomly scattered secondary tracks would give a big worsening in the jet algorithm performance as the information on primary particle is completely lost. At this stage, it is thus very important to remove as much as possible the secondary tracks from the Event, even if this can involve some loss of information. The selection criteria developed in this work, which allow to retain only primary tracks, consist in a set of cuts on the reconstructed track $\eta$, $Z_0$ and reduced $\chi^2$. More precisely, the reconstructed parameters of a primary track have to satisfy the following conditions:

- the reconstructed $\eta$ should be compatible with the nominal acceptance of T2 ($5.3 < |\eta| < 6.5$);
- the reconstructed $Z_0$ should be compatible with the $Z_0$ of a track coming from the IP; detailed studies on this cut is presented later in this section;
- the track $\chi^2$ should be less than 1; this condition has been derived by comparing the
left and the right plot of fig. 3.9, from where it is clear that all the primary tracks are characterized by $\chi^2 < 1$.

The cut on track $Z_0$ requires in particular a more detailed study. This parameter is expected to be $\eta$ dependent since the amount of beam pipe material crossed by the particle depends on its pseudo-rapidity. In addition, as detailed in next subsection, for a given intrinsic resolution on the reconstructed polar angle ($\Delta \theta$) the $\eta$ resolution ($\Delta \eta$) increases with $\eta$ itself, so that the $Z_0$ of the reconstructed track spans a largest $Z$ interval across the IP as the track reconstructed $\eta$ increases. Furthermore, once the track $\eta$ is fixed, multiple scattering is expected to have more influence for lower energy particle, as shown in formula 1.10 from where it is clear that the lower is the particle energy, the higher is the RMS of the multiple scattering angle distribution ($\langle \theta \rangle$), so the higher is the interval spanned by the $Z_0$ parameter. Therefore the cuts on the reconstructed track $Z_0$ should depend on the track reconstructed $\eta$ and in principle also on the track energy (even if the latter information is not available from T2 data\textsuperscript{12}).

In order to understand the dependence of the primary track $Z_0$ distribution on track energy and pseudo-rapidity, 43000 and 50000 $\pi^-$ with energy of 50 GeV and 10 GeV respectively have been generated with a flat azimuthal distribution in $0 \leq \phi \leq 2\pi$ and a flat pseudo-rapidity distribution in the interval $5.2 \leq \eta \leq 6.5$. This pseudo-rapidity interval has then been divided into sub-intervals of $\Delta \eta = 0.02$ and, for each of them, the $Z_0$ distribution of the reconstructed tracks having their $\eta$ falling in that sub-intervals, has been found. These distributions have then been fitted with a gaussian and its standard deviation $\sigma_{\Delta Z}$ has been computed. Some examples of the $Z_0$ distribution, obtained with 50 GeV $\pi^-$, in four different pseudo-rapidity sub-intervals, are shown in fig. 3.10. From this figure it is clear that, if the reconstructed track $\eta$ is around 5.53 or at the edges of the T2 acceptance such fitting cannot be realized because the distribution are very different from a gaussian, so a different procedure has to be adopted in order to derive a proper $Z_0$ cut in these zones.

Excluding these cases, the values of $\sigma_{\Delta Z}$ have been plotted versus $\eta$ so to derive a functional relation (where $\eta$ is now the center of the pseudo-rapidity sub-interval). In

\textsuperscript{12}This information can in principle be derived from CASTOR calorimeter, when a common CMS/TOTEM data taking will be possible.
the pseudo-rapidity regions where $\sigma_{\Delta Z}$ is not available an entry equal to the closer well-reconstructed $\sigma_{\Delta Z}$ is considered. An analytical expression which describes $\sigma_{\Delta Z}(\eta)$ has then been found via a polynomial fitting in different pseudo-rapidity intervals. Such curve, obtained for 50 GeV $\pi^-$, is shown in fig. 3.11 (bottom). In order to understand the dependence of the $\sigma_{\Delta Z}(\eta)$ curve on the energy, the same study has been also performed for 10 GeV $\pi^-$; the corresponding curve is shown on top of fig. 3.11.

These curves are important because they are used to derive the cut on reconstructed track $Z_0$. More precisely, given a track with reconstructed parameters $\eta$ and $Z_0$, it is
3.5 Tracking performance

![Graphs showing \( \sigma_{\Delta Z}(\eta) \) for tracks with \( \chi < 1 \) in T2, with different energies (10 GeV and 50 GeV). Each point represents the standard deviation of the Gaussian fit for the \( Z_0 \) distribution related to reconstructed tracks having \( \eta \) inside a pseudo-rapidity sub-interval centred at the value shown in the x-axis. The value around \( \eta = 5.53 \) and at the edge of the T2 acceptance are set equal to the closer reconstructed points.]

Figure 3.11: \( \sigma_{\Delta Z}(\eta) \) curves as obtained with 10 GeV \( \pi^- \) (top) and 50 GeV \( \pi^- \) (bottom). Each point represents the standard deviation of the Gaussian fit for the \( Z_0 \) distribution related to reconstructed tracks having \( \eta \) inside a pseudo-rapidity sub-interval centred at the value shown in the x-axis. The value around \( \eta = 5.53 \) and at the edge of the T2 acceptance are set equal to the closer reconstructed points.

Considered a primary track if the following condition is satisfied:

\[
Z_0 < 3 \sigma_{\Delta Z}(\eta)
\]  

(3.15)

It is important to remark that it is not possible to reconstruct the energy of the particle associated to the track, so in principle the proper energy-dependent curve \( \sigma_{\Delta Z}(\eta) \), used for the primary track selection, is unknown.
As shown in fig. 3.11, for a fixed track \( \eta \), the difference between the two curves taken as reference is around 300 ÷ 400 mm (which means a not negligible 40 ÷ 50\%). Therefore, by choosing the curve obtained with pions of higher energy, a more strict selection on the track \( Z_0 \) is obtained, which reflects in an enhanced possibility for primary low energy tracks rejection. For what concerns the T2 telescope, 10 GeV pions can be considered quite low energy particles, when compared to the mean charged particle energy expected in the region 5.3 < \( \eta \) < 6.5 for typical inelastic processes. In fig. 3.12, the charged particle energy distribution for minimum bias events is shown for particles falling in the T1 (left) and T2 (right) acceptance. Therefore, while for T2 it is not expected a big fraction of particles with energy below 10 GeV, in T1 the mean charged particle energy is sensibly lower, so that also the reconstruction of particles having energy of few GeV can be important. The dependence on the energy of the curves in fig. 3.11 is a feature that will be used in tracks selection for topological jet algorithms (see chapter 4) where it is desirable to have a high selection efficiency only for high energy tracks. Fig. 3.13 shows the charged particles energy distribution for Di-jet event (left) and Single Diffractive event (right) in the T2 acceptance region. As the mean particle energy of typical hadronic soft processes...
3.5 Tracking performance

![Charged Particle Energy in T2 (di-jet event)](image)

![Charged Particle Energy in T2 (SD event)](image)

Figure 3.13: Charged particle energy distribution for Di-jet events (right) and for Single Diffractive events (left), in the T2 acceptance region.

is lower than the mean particle energy of processes involving very forward jets, a relative bigger rejection of tracks in background events is expected than in Di-jet events, when (for instance) the bottom plot curve of fig.3.11 is considered in order to apply the $Z_0$ cut according to eq. 3.15. Fig. 3.14, obtained for 50 GeV $\pi^-$, shows the track reconstruction pseudo-efficiency plotted as a function of the generated particle $\eta$, as obtained by applying different cut combinations on reconstructed track $\chi^2$, $\eta$ and $Z_0$. The top-left plot is the same picture already reported in fig. 3.7, where no track selection criterion is applied; the top-right picture is obtained by considering only the reconstructed tracks with $\chi^2 < 1$, which gives a first sensible reduction in secondary track reconstruction while still allowing a very high primary track reconstruction pseudo-efficiency; the central-left picture shows that the most relevant cut is introduced by requiring that the track $Z_0$ satisfy the eq. 3.15 (the curve used to evaluate $\sigma_{\Delta Z}(\eta)$ is from bottom plot of fig. 3.11, obtained with 50 GeV $\pi^-$); the central-right picture is obtained by using track with reconstructed pseudo-rapidity in the range $5.2 < \eta < 6.5$; in the bottom-left picture the cuts on track $\eta$ and $Z_0$ are used together; finally, the bottom-right picture is obtained by requiring all the three cuts on reconstructed $\eta$, $\chi^2$ and $Z_0$ (hereafter, this combination in called “standard cuts”).

As expected, the additional request of reconstructed tracks $\eta$ in the range $5.2 < \eta < 6.5$
3.5 Tracking performance

Figure 3.14: Mean number of reconstructed tracks for a 50 GeV $\pi^-$ generated at the pseudo-rapidity shown in the x-axis and with a flat azimuthal distribution. The effects of different cut combinations used in order to select primary tracks are shown.

has a little effect if the cut on $Z_0$ is already performed (which means these cuts are strongly correlated). As shown in the bottom-right picture of fig. 3.14, a single track reconstruction pseudo-efficiency well above 80% is obtained in all the T2 acceptance even if the full set of standard cuts is applied (excluding, of course, the zone around $\eta = 5.53$ where the
primary particle interaction with the beam pipe is very probable). However a better pseudo-efficiency for $\eta > 6.1$ can in principle be achieved. Indeed, at high $\eta$ the curve in fig. 3.11 as obtained with 50 GeV pions should be corrected because the $Z_0$ distributions in that zone are no longer well fitted by a gaussian. More detailed studies should allow the use of more conservative (larger) values of $\sigma_{\Delta Z}(\eta)$ to be adopted at high $\eta$, with a consequent increase in primary track reconstruction pseudo-efficiency. Moreover, also the cut on reconstructed track $\eta$ which is at the moment considered may be too strict. The multiple scattering on high-$\eta$ particles (which are anyway detected by T2) can lead to a reconstructed particle pseudo-rapidity which is outside the nominal T2 acceptance.

### 3.5.3 Angular resolution on track reconstruction

Some results about the angular resolution characterizing the track reconstruction in T2 are reported in this section. From simulated events it is possible to study the difference between the reconstructed track $\eta$ and $\phi$ parameters and the $\eta$ and $\phi$ coordinates of the generated particle. This difference is called “residual” and the term “resolution” is here used for the standard deviation of the gaussian fit for the distribution of these residuals. The angular resolution for track reconstruction in T2 is a quantity which is $\eta$ dependent and energy dependent. This dependence is due to the particle-beam pipe interaction and to magnetic field effects. In the forward region, the difference $\Delta \theta$ between the polar angles of two close tracks (or the difference between the generated particle $\theta$ and the reconstructed track $\theta$), corresponds to a difference in $\Delta \eta$ given by:

$$
\Delta \eta = e^{\vert \eta \vert} \left( 1 + e^{-2\vert \eta \vert} \right) \Delta \theta
$$

(3.16)

Therefore, $\Delta \eta$ depends on $\Delta \theta$ and $\eta$, increasing with $\eta$.

Because of the presence of the CMS magnetic field in the central zone (around the IP), the azimuthal angular resolution is a quantity which strongly depends on the particle energy. Indeed the variation of the charged particle azimuthal angle, due to particle curvature in the magnetic field, is inversely proportional to the particle energy itself: $\Delta \phi \propto 1/E$. Therefore, it is only possible to define an “intrinsic” $\phi$ resolution, i.e. the azimuthal resolution in absence of magnetic field effects. Fig. 3.15 shows the pseudo-
rapidity (top) and the azimuthal (bottom) residual distribution as obtained by firing 50 GeV $\pi^-$ at $\eta = 5.9$ (left) and $\eta = 5.59$ (right). The simulation at $\eta = 5.59$ shows the effects on track resolution due to particle interaction with the beam pipe, while at $\eta = 5.9$ this effects are not relevant. The reconstructed tracks which have been considered in order to obtain these distributions are selected by using the standard cuts presented in section 3.5.2\textsuperscript{13}. As already stated, because of particle $Z$ smearing around the IP at generation, it is possible that a particle generated very close to $\eta = 5.53$ will actually cross not a big amount of beam pipe material. As shown in the bottom plots, because of the magnetic field, 50 GeV charged pions suffer a systematic shift in reconstructed track $\phi$ of about $4^\circ$ with respect to the generated particle azimuthal angle.

The intrinsic pseudo-rapidity and azimuthal resolution for 50 GeV $\pi^-$, as a function of

\textsuperscript{13}Some contribution due to fake tracks from secondary interactions is anyway expected, also after applying the standard selection cuts, in all distribution obtained for particle $\eta$ around 5.53.
the generated pion \( \eta \), are shown in fig. 3.16. Each point of these plots is obtained from a gaussian fit on a histogram of the kind shown in fig. 3.15. Some points near \( \eta = 5.53 \) are not included because at these pseudorapidities the distributions of the residuals have a number of entries which is only 5%-10% of the total number of events generated so that it is not possible to derive an estimation of the resolution from a gaussian fit. Fig. 3.16 (left) shows that the pseudo-rapidity resolution gets worse as the particle pseudo-rapidity increases, which is consistent with equation 3.16. For 50 GeV pions, the pseudo-rapidity resolution ranges from 0.04 to 0.07, excluding the critical points near \( \eta = 5.53 \), while an intrinsic azimuthal resolution of about 1° is obtained, compatible with the value \( \Delta \phi_{PAD}/\sqrt{12} \) which can be taken as an upper limit on the azimuthal resolution, where \( \Delta \phi_{PAD} = 3^\circ \) is the pad azimuthal angle. Fig. 3.17 shows the pseudo-rapidity and azimuthal angle residual distribution as obtained by firing 50 GeV \( \pi^- \) (right) and 10 GeV \( \pi^- \) (left) uniformly in the pseudo-rapidity and azimuthal acceptance of T2. As expected, due to multiple scattering effect, the intrinsic pseudo-rapidity and azimuthal resolution get worse as the pion energy decreases: the FWHM of the distributions get larger roughly by a factor 2 when going from 50 to 10 GeV energy. However, even for 10 GeV pions, which are expected to represent the low energy hadronic component for particles produced in typical inelastic processes and detected in T2, the average intrinsic pseudo-rapidity resolution is below 0.1 although, at
3.5 Tracking performance

This energy, the azimuthal shift due to the magnetic field is of about 20°. The systematic shift of the mean value of the pseudo-rapidity residuals distribution shown in fig. 3.17 (top), which is larger as the particle energy decreases, can be due to non parallel (respect to the beam axis) components of the CMS magnetic field that can change, especially for low energy particle, its polar angle.

3.5.4 Reconstruction of tracks from neutral pion decay

In this section the reconstruction of tracks related to charged particles originated by a primary π^{0} is investigated. The dominant decay of a neutral pion is the electromagnetic decay into two photons. The subsequent photon conversion into e^{+}e^{-} pairs, which can be detected with T2, can happen because of the photon interaction with the beam pipe steel or with the material outside the beam pipe itself (for instance with the CMS HF calorimeter, which has an acceptance of 3 < |η| < 5). Occasionally, when a large amount
3.5 Tracking performance

of material is crossed, also an electromagnetic shower can occur with the production of many charged particles ($e^+, e^-$).

In the center of mass (CM) frame of the $\pi^0$, the opening angle of the two photons is obviously 180°. The laboratory frame is here defined as the frame at rest with the IP, having its x axis along the pion trajectory before the decay. If $\theta_0$ is the emission angle of one of the two photons in the CM frame, the opening angle between the two photons in the laboratory frame ($\Theta$) is given by the following relation [43]:

$$\cos \Theta = \frac{2V^2 - 1 - V^2 \cos^2 \theta_0}{1 - V^2 \cos^2 \theta_0}$$

where $V$ is the relative velocity between the CM and the laboratory frame. Putting inside eq. 3.17 the velocity of a 50 GeV $\pi^0$, it is found that all decays with $0.7 \text{ rad} < \theta_0 < \pi/2 \text{ rad}$ give an opening angle in the laboratory frame of $5 \text{ mrad} < \Theta < 7 \text{ mrad}$ while the opening angle is bigger\(^{14}\) for $0 \text{ rad} < \theta_0 < 0.7 \text{ rad}$. This means that when $\pi_0$ is generated at a given polar angle $\theta_i$, its decay photons will most probably emerge from IP at $\theta_i \pm \Delta \theta$ where, very roughly, $\Delta \theta \leq 2\pm 3 \text{ mrad}$. Consequently, we can expect that a $\pi_0$ originated outside the T2 detector acceptance can contribute charged tracks in T2, which are associated to $e^+ - e^-$ produced in secondary interactions of decay photons. Fig. 3.18 shows the mean number of reconstructed tracks in T2 for a 50 GeV $\pi^0$ as a function of its pseudorapidity.

In this figure, the track reconstruction is performed without (left) and with (right) the beam pipe simulation, applying (bottom) or not (top) the standard cuts for primary track selection. Similar plots for 50 GeV $\pi^-$ are also shown in fig. 3.19. There are several important things to note when comparing the plots of fig. 3.18:

- around $\eta = 5.2 - 5.3$, the mean number of reconstructed tracks has a maximum also if the beam pipe is not simulated (see top-left plot); this can be explained with the interaction of the photons in the lower part of the HF calorimeter (see fig. 1.12);

- the dominant effect is given by the beam pipe: adding its simulation the average track multiplicity increases by about a factor of 6, in particular around $\eta = 5.2 -$

\(^{14}\)An isotropic decay is defined by $\frac{dN}{d\Omega_0} = C$, where $C$ is a constant and $\Omega_0$ is the solid angle in the CM frame. Assuming azimuthal symmetry, $\frac{dN}{d\eta} \sim \sin \theta_0$. Therefore, the bigger fraction of the decays happens in the range $0.7 \text{ rad} < \theta_0 < \pi/2 \text{ rad}$. 
3.5 Tracking performance

Figure 3.18: Profile plots obtained firing 50 GeV neutral pions uniformly with $4.5 < \eta < 7.5$ and $0 < \phi < 2\pi$ rad without (left) and with (right) the simulation of the beam pipe. The track reconstruction is obtained by applying (bottom) or not (top) the cuts on $\chi^2$, $\eta$ and $Z_0$ for primary track selection.

5.3 (see top-right plot); in this pseudo-rapidity range a photon from $\pi^0$ decay can go in the region between the beam pipe and the lower part of the HF calorimeter, convert into an $e^+ - e^-$ pair or start an electromagnetic shower in the latter and then contribute several particles in T2 due to further interactions in the beam pipe; this happens because secondary particles generated in the lower part of HF and travelling towards T2 have to cross in average several cm of beam pipe steel before reaching it;

- including the beampipe, a rising of the mean number of reconstructed tracks is also found when $\pi^0$ are generated around $\eta = 7$ (see the top-right plot); as already discussed in chapter 1.3, this behaviour is expected because of the interactions with the cylindrical part of the beam pipe upon which T2 is placed; however, it is important to note that the increase in charge multiplicity around $\eta = 7$ obtained with $\pi^0$...
shown in fig. 3.18 (top-right), is less sharp when compared with the correspondent plot in fig. 3.19. This is attributed to the effect of the non negligible opening angle between the two photons. Indeed, it is important to remind that the x-axis reports the pseudo-rapidity of the generated $\pi^0$ and not of the photons (which will actually generate the charged secondaries). At high $\eta$ and without the beam pipe, the photons have a higher probability to cross T2 without interacting before with other materials, which explains the decreasing trend at high $\eta$ shown in the top-left plot of fig 3.18;

- finally, it is important to notice that, even if the selection cuts allow to reject a big fraction of secondary tracks (fig. 3.18, bottom), $\pi^0$ decays contribute to increase fake track reconstruction in the event (i.e. tracks from secondary interactions which

Figure 3.19: Profile plots obtained firing 50 GeV charged pions uniformly with $4.5 < \eta < 7.5$ and $0 < \phi < 2\pi$ rad without (left) and with (right) the simulation of the beam pipe. The track reconstruction is obtained by applying (bottom) or not (top) the cuts on $\chi^2$, $\eta$ and $Z_0$ for primary track selection.
satisfy all selection cuts). As consequence of that, the selection of primary tracks in a typical physics event where also neutral pions are present, can become more difficult with a resulting loss in primary track reconstruction efficiency.
Chapter 4

Topological jet algorithms

As already reported in chapter 2, several inelastic processes are expected to produce jets in the T1/T2 pseudorapidity acceptance. All these processes can be of interest for the TOTEM physics programme. However, in order to perform such kind of studies within the TOTEM experiment, jet reconstruction with T1/T2 detector tracks is required. As the TOTEM inelastic telescopes are outside the CMS magnetic field, this represents by itself a challenge as jet algorithms are usually based on the clusterization of either particle energy deposition into calorimeters or charged track transverse momenta as derived from deflection into magnetic field. In this chapter two innovative jet algorithms, which have been developed in order to allow jet reconstruction only relying on charged particle topological information, are presented. The interest on jet reconstruction at track level is motivated in this part of the thesis work also by the fact it represents a powerful tool in order to check the T2 track algorithm performance in a high track multiplicity event topology. These jet algorithms, inspired by traditional $k_t$ and Cone algorithms, were developed and tested at the particle level on simulated Pythia Di-jet and Single Diffractive events. The former sample has been used to understand the performance of the algorithms, the latter has been considered in order to have an idea of how much background the algorithms pick up in terms of fake jet reconstruction. In these studies the algorithms were developed with attention on the results obtained looking in the region covered by the TOTEM inelastic telescopes T1 ($3.1 < \eta < 4.8$) and T2 ($5.3 < \eta < 6.5$). Anyway it should be pointed out that the algorithms implementation is in principle expected to be valid in any $\eta$ region.
4.1 Jets and jet algorithms

After a brief overview about jets (section 4.1) and a description of the MC samples used (section 4.2), a detailed description of these “topological” jet algorithms is given in section 4.3. In section 4.4 the parameters characterizing the algorithms and the jet selection criteria will be assigned according to simulation results. In section 4.5 the studies done at charged particle level, which allow to understand the effective performance of the algorithms without any bias due to track reconstruction efficiency, magnetic field deflection or additional tracks from secondary interactions, are presented. The same studies have then been performed by using the jet algorithms on T1 and T2 reconstructed tracks. Track selection criteria, as derived in chapter 3, have been applied at this stage in order to allow an optimal algorithm performance. However, it is important to notice that the algorithm implementation inside the CMSSW framework allows for potential improvements in performance which can in principle derive from future detailed studies, eventually performed on specific physics events of interest. It should also be pointed out that these algorithms can in principle be combined with the CMS HF/CASTOR calorimeters information for an optimized jet reconstruction in the very forward region, given the limited azimuthal and the lack of polar segmentation of the CASTOR calorimeter. Future studies are foreseen in this direction.

4.1 Jets and jet algorithms

Most of the pp hadronic interactions are “soft”, with a small momentum transfer between the colliding partons. Occasionally, head-on collisions between the incoming hadrons, in which partons interact with large momentum transfer, occur. These “hard” processes are characterized by a colored system of outgoing partons with high \( p_T \) which then fragments in groups of spatially collimated energetic particles. These energy flows are called “jets”. The studies of such kind of events have always been considered important, first of all because the total four-momentum of the final jet particles is equal to the four-momentum of the original parton. Therefore, the reconstruction of jet properties gives access to the parton dynamics and allows to test the predictions of perturbative Quantum Chromo Dynamics (pQCD), which is the most important predictive theory for hard hadronic interactions. Fig. 4.1 shows the production of jets, as simulated in hadronic scattering. The partons involved
4.1 Jets and jet algorithms

Figure 4.1: Illustration of the model used in QCD Monte Carlo to simulate a proton-proton collision in which a hard $2 \rightarrow 2$ scattering with transverse momentum $p_t$ has occurred [44]. The resulting event contains particles that originate from the two outgoing partons plus initial and final state radiation and particles that come from the breakup of the protons (“beam-beam remnants”).

in the jet production can emit gluon radiation prior to the short distance scattering (this step is called Initial State Radiation or ISR). The remnants of the original hadrons are no longer in a color singlet state and soft interaction can take place between them. The hard scattering between the selected partons is simulated at leading order or next to leading order in pQCD [22].

The outgoing hard partons also emit Final State Radiation (FSR): the FSR, also simulated in Monte Carlo programs, consists in the emission of soft gluons dominated by collinear emission respect to the original parton direction. The last step of jet production consists in the fragmentation (also called hadronization) of all outgoing partons in the final state hadrons. The hadrons coming from all partons which don’t participate in the hard scattering contribute to the so called “underlying event” (UE) [45]. Jet algorithms are used to eventually find and reconstruct jets which occur in the event. Traditional jet algorithms (when run on simulated events) cluster partons or particles or calorimeter towers according to their proximity in coordinate space (as for example in cone algorithms) or proximity in momentum space (as for example in $k_T$ algorithms). Obviously, when run on real data events, these algorithms can only performs a clusterization at calorimeter or reconstructed track level. More details about these jet algorithms will be given in section 4.3.1 and 4.3.2; instead, here it is important to remind some basic features which
are commonly required to any jet algorithm. For a precise comparison of experiment to theory, a jet algorithm should give very similar results when it is applied to a state with just few outgoing partons (as in NLO perturbative theory), to a state with many partons (after the Monte Carlo simulation of the parton showering), to a state with hadrons (as simulated in a Monte Carlo which includes models of the hadronization and for the UE) to the observed tracks or energy deposition in a real detector. The algorithm should be also infrared and collinear safe. The formal definition of infrared safety is the following [46]:

- an observable is infrared safe if, for any n-parton configuration, adding an infinitely soft parton does not affect the observable at all;

while the formal definition of collinear safety is:

- an observable is collinear safe if, for any n-parton configuration, replacing any mass-less parton by an exactly collinear pair of mass-less partons does not affect the observable at all;

Should the above requirements be unsatisfied, there would not be correspondence between the parton energy and the reconstructed jet energy. In analytical pQCD calculations, soft parton emission gives a divergent contribution to the cross section of the jet event which is cancelled by the collinear emission contribution: in this sense, the theory is infrared and collinear safe. This independence on soft/collinear emission, should be maintained also by the jet algorithms. In section 4.3 the issue of infrared and collinear sensitivity in cone-based jet algorithms is investigated into more details.

### 4.2 Monte Carlo samples

Pythia6 is one of the most important Monte Carlo tools used by the CMS Collaboration for event generation [35]. Pythia is a general purpose event generator containing theory and models for a number of physics aspects, including hard and soft interactions, parton distributions, multiple interactions, initial and final state radiation, parton showers, fragmentation (“Lund” model) and decays. For the present studies at particle level (see subsection 4.5.1) two samples of 5000 events have been generated for Single Diffractive and Di-jet events using Pythia6.409, which is the standard version currently used in the release
1.75 of CMSSW adopted for this work. The studies at charged track level (see subsection 4.5.2) are indeed performed on 900 Di-jet events and 640 Single Diffractive events which also includes the simulation of the T1 and T2 detectors.

### 4.2 Monte Carlo samples

#### 4.2.1 Di-jet events

One analysis of interest which motivates the current study is the proton dissociation into 3 jets. This process (see subsection 2.4.2) only involves forward quark-jets coming from proton dissociation without initial state radiation and underlying event effects. As a dedicated Monte Carlo for the simulation of this process is still under validation [25], the jet algorithms are developed and tested in two jets (Di-jet) events originated by quarks only. Di-jet events are generated configuring Pythia to use the following fundamental hard processes:

\[
\begin{align*}
q q & \rightarrow q q \\
qu q' & \rightarrow q q' \\
q \bar{q} & \rightarrow q \bar{q} \\
q \bar{q} & \rightarrow q' \bar{q}' \\
q' \bar{q}' & \rightarrow q \bar{q}
\end{align*}
\]

where both partons are generated with \(5.0 < \eta < 6.5\) and \(p_t > 3\) GeV/c. Moreover, multiple parton interaction and QCD/QED initial state radiation are switched off. Looking carefully at the generated events, it is found that, during fragmentation, almost 50% of the events show interactions among the strings produced by the initial outgoing partons and the beam remnant particles. This causes a large multiplicity in the final state and it is chosen to reject this type of events because the final particle distribution is quite different respect to the two jets shape useful to test the jet algorithm performance in a situation similar to the one expected in the reconstruction of the events from the proton dissociation. The invariant mass of the two outgoing partons is evaluated and compared to the invariant mass of the final particles which are daughters of them. A difference between these invariant masses, being the signature of the unwanted interactions, is used for the selection of two-jet events without jets/beam-remnant interaction.
4.2.2 Single Diffractive events

For diffractive and non diffractive soft events, Pythia implements some phenomenological models to describe these processes [47]. Single Diffractive events

\[ p + p \rightarrow p + X \]

are generated by requiring the unfragmented proton to fall every time in the same arm of the TOTEM Roman Pots detectors (with \( \eta < 0 \)).

4.3 Topological jet algorithms

The basic feature of both jet algorithms developed in this thesis and presented in this section, is the exchange of particle momentum information with the density of the detected particles in \( \eta - \phi \) space. In fact, the algorithms will be used on T1/T2 data, so they cannot make use of calorimetric information or trajectory measurement inside magnetic field. As the T1/T2 telescopes are sensitive only to charged particles, in what follows it is assumed that the algorithms have already selected the stable charged particle list falling in the T1/T2 \( \eta \) region. In these studies the two different jet algorithms are developed and tested by using the same physics event file. The first is a cone-like algorithm, improved with “MidPoint” and “Splitting and Merging” procedures, inspired to the MidPoint cone jet algorithm successfully used at Tevatron in Run2 [48]; the second is inspired to a simplified version of the \( k_t \) algorithm, as introduced in reference [49].

4.3.1 Cone algorithm

In this section the topological cone algorithm will be described. In what follows, the use of certain numerical values will be justified in section 4.3.3. The algorithm implementation can be summarized in the following steps (also represented in fig. 4.2):

- build an \( \eta - \phi \) grid, with cell dimension (\( L \)) not too small in order to save computation time and not too big in order to avoid resolution reduction in the reconstructed jets;
• compute the number of particles in every cell and build an ordered list according to the number of particles inside the cells. Doing this, exclude the neighbour cells of a seed, i.e. a cell with a major number of particles. If the cell size is too small $L \sim 0.3$, exclude all the cells inside a square with side $R \sqrt{2}$ centred on the seed;

• for every remaining cell in the ordered list start the search of particles in a cone with radius $R$ around the cell centroid, compute the new centroid as the mean of the particles position falling inside, repeat the procedure until the cone is stable (i.e. the list of the particles falling inside the cone remain unchanged between two subsequent iteration);

• at this point there is a list of candidate jets. In order to decrease “infrared sensitivity” and some “collinear sensitivity” (see section 4.3.1) start the “MidPoint” procedure which adds another cone in the middle of every pair of cones with distance less than $2R$;
• the jets in the list can have some particles shared between two or more jets. Use the “Splitting and Merging” procedure (see section 4.3.1) to univocally assign every particle to only one jet;

• refine the jet-list including some cuts on the jet particle multiplicity and the jet shape.

More details on the MidPoint and Splitting and Merging procedures are given in the following two subsections.

The MidPoint procedure

A traditional cone algorithm (which uses energy information instead of particle density) is affected by “infrared” and “collinear sensitivity” [48]. The “MidPoint” procedure is a tool to reduce this type of problems. Generally, if we require that jet algorithm results are insensitive (i.e. give the same list of reconstructed jet) to the presence of soft radiations around the real jets, the situation shown in fig. 4.3 should be avoided. Fig. 4.3 shows

Figure 4.3: The issue of infrared sensitivity: the picture shows how the presence of soft radiation between two jets may cause a merging of the jets that would not occur in the absence of the soft radiation.

the jets (represented with cones) reconstructed by a traditional cone algorithm whether soft radiation (represented as wavy arrows) between the hard partons (represented as straight arrows with length proportional to the seed energy) is present or not. Without soft radiation the algorithm reconstruct two jets (fig. 4.3, left). If soft radiation is present between two jets (fig. 4.3, right) and the jets distance is less than $2R$, the two jets will be probably merged. Introducing the MidPoint procedure, the algorithm is forced to look for a new jet between every pair of jets previously found. Therefore in this case the jet
4.3 Topological jet algorithms

algorithm results, whether or not infrared radiation is present, are the same. Even if this method can seem to strict (for example, real two jet event can be identified as a single jet) it is important to remind that one of the most important feature for a jet algorithm is the infrared safeness. If this requirement would not be satisfied, it would not be possible to quantify the contribution of real two jet reconstructed as a single one, and the prediction power of the PQCD could not be tested on the data. For the topological cone algorithm developed in this work, the MidPoint procedure is useful to account for missing $\pi^0$ in the jets cores: for instance, consider the case with two charged particles separated by more than $R$ in the $\eta - \phi$ plane. It is possible that these particles belong to the same jet; however the cone algorithm without MidPoint would not find any jet if these particles belongs to different seeds, because the effect of the multiplicity cut (assumed bigger than one) which is applied on the final jet list. A jet is instead found if the algorithm would work even with potential $\pi^0$ in the core of the jet or if the charged particles mentioned above would belong to the same seed. This “infrared-like” sensitivity is removed by the using of the MidPoint procedure in the cone topological algorithm, which force the algorithm to look in a cone which embed both the charged particles.

Another issue present in traditional cone algorithms is the dependence of the algorithm results from the jet energy sharing between calorimetric towers in the $\eta - \phi$ plane: the term “collinear sensitivity” is used for this issue. A similar issue can also affect a density-based jet algorithm and fig. 4.4 can be useful in order to understand the problem. The picture in

![Figure 4.4: The issue of collinear sensitivity: the splitting of particles belonging to the same jet among several grid cells could lead to the absence of the seed or could produce a different ordering (in particle number) of the seeds.](image)
fig. 4.4 (left) shown an event where three seeds (for the cone algorithm here presented, the length of the arrow is not proportional to the seed energy but to the number of particles inside the seed) are found; the algorithm starts from central most populated cell and it will probably find 3 jets or only 1 jet. The picture in fig. 4.4 (right) is the same process where the particles are shared between two cells, for instance because of a relative shift of the grid position. In this case, without MidPoint, the algorithm starts from the right-most cell and it will find only 2 jets. With Mid Point it will find 3 jets o 1 jet as before, thus removing this “collinear-like sensitivity”.

The Splitting and Merging procedure

Splitting and Merging procedure must be implemented in order to avoid particle sharing between two or more jets (see fig. 4.5) so that in the final jet list every particle is univocally assigned at most to one jet. In the common implementation of the jet algorithm, the input parameter “$f_{\text{merge}}$”, which control the assignment of the shared particle, has to be set between 0 and 1. For the topological cone algorithm $f_{\text{merge}} = 0.5$ is chosen which imply that if the particles shared between two jet are less than 50% of the particles forming the less populated jet of the two, the algorithm will assign the particles to the jet having the closer centre and it will remove these particles from the other; otherwise if the shared
particles are more than 50%, the two jet will be merged in a single jet, removing the initial two overlapping jets from the list.

### 4.3.2 $k_t$ algorithm

Before the description of the topological $k_t$ algorithm developed in this work, it is useful to recall the implementation of a traditional $k_t$ algorithm, as described in [48].

The $k_T$ jet algorithm starts with a list of preclusters which are formed from calorimeter cells. For each precluster, a vector

\[ (E, p) = E(1, \cos \phi \sin \theta, \sin \phi \sin \theta, \cos \theta) \]

is assigned where $E$ is the energy associated with the precluster, $\phi$ is the azimuthal angle, and $\theta$ is the polar angle with respect to the beam axis. Then, the square of the transverse momentum $k_T^2$, and the pseudorapidity $\eta$ are calculated for each precluster. The algorithm steps are as follows:

- For each precluster $i$ in the list, define $d_i = k_T^2; i$; for each pair $(i, j)$ of preclusters $(i \neq j)$, define
  \[ d_{ij} = \min(k_T^2, k_T^2) \frac{\Delta R_{ij}^2}{D^2} \]  
  (4.1)
  where $\Delta R_{ij}^2 = (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2$ and $D^2 \sim 1$ is a parameter of the algorithm. For $D = 1$ and $\Delta R_{ij} << 1$, $d_{ij}$ is the minimal relative transverse momentum squared of one vector with respect to the other;

- find the minimum of all the $d_i$ and $d_{ij}$ and label it $d_{\text{min}}$;

- if $d_{\text{min}}$ is a $d_{ij}$, remove preclusters $i$ and $j$ from the list and replace them with a new, merged precluster described by $(E_i + E_j, p_i + p_j)$;

- if $d_{\text{min}}$ is a $d_i$, remove the precluster $i$ from the list of preclusters and add it to the jet list;

- if any preclusters remain, go to the first step.
The lack of seeds make the $k_t$ jet algorithm free from infrared and collinear sensitivity. The $k_t$ algorithm developed in this work is based on the Cambridge/Aachen algorithm described in [49] where it is also rigorously proved the equivalence (for what concern the particles clusterization to make a jet) in the exchange of the particle-distance information with quantities involving their transverse momentum as commonly used in the traditional $k_t$ jet algorithms.

In fact, two basic facts, as reported below, should be noted in order to “translate” part of the traditional $k_t$ algorithm already described, in a topological algorithm:

1. The minimum distance $d_{ij} = \min(k_{ti}^2, k_{tj}^2)R_{ij}^2$ is obtained for the pair of particles with the smallest $R_{ij}$, i.e. with the minimal distance in the $\eta$ - $\phi$ plane. Indeed, suppose this is wrong and that there exists a particle $l$ such that $R_{il} \leq R_{ij}$: then $d_{il} = \min(k_{ti}^2, k_{tl}^2)R_{il}^2$ and since $\min(k_{ti}^2, k_{tl}^2) \leq k_{ti}^2$ it would result $d_{il} \leq d_{ij}$ in contradiction with the statement that $i$ and $j$ have the smallest $d_{ij}$.

2. During the operation of a traditional $k_t$ algorithm the minimum between all the $d_i$ and $d_{ij}$ has to be found. Suppose that the minimum is given by a $d_i$: than $k_{ti}^2 < \min(k_{ti}^2, k_{tl}^2)R_{ij}^2 \forall l,j$. Because $k_{ti}^2$ is the minimum $k_{ti}^2 < k_{ti}^2R_{ij}^2$; therefore, if $d_i$ is a minimum, it should happen $1 < R_{ij}^2 \forall l,j$ and $d_i < d_k \forall k$. Now suppose instead that the minimum is given by the distance $d_{in}$ of a pair of particles $i, n$. This time, $\forall l, \min(k_{ti}^2, k_{tn}^2)R_{in}^2 < k_{tn}^2 \Rightarrow R_{in}^2 < 1$. Therefore, in order to obtain the minimum element of the set $\{d_i, d_{ij}\}$, the minimum $R_{ij}$ has to be found. If $R_{ij} < 1$ the particles to merge are the particle $i, j$, otherwise the minimum in the $\{d_i, d_{ij}\}$ is given by the minimum $d_i$.

Now it is possible to summarize the implementation of the topological $k_t$ jet algorithm developed in this work:

- build the list of the precluster: at the beginning, each precluster is made by only one particle;

- for each precluster $i$ establish its nearest neighbour, that is find the precluster $j$ with minimum distance $R_{ij}$. Find the minimum of the $R_{ij}$;
• if the minimum $R_{ij}$ is less than one, remove $i, j$ from the list and merge them in a new jet, otherwise $i$ is a jet and it’s not considered anymore. It should be noted that, being added to the jet list, in the routine $i$ could represent a set of many particles;

• if $i$ is not at the end-list, restart from the first step, otherwise the procedure is completed.

In order to summarize the above steps, the algorithm is also represented in fig. 4.6. It should be pointed out that the particle clustering condition is equivalent to those of the traditional $k_t$ algorithm; however the merging of the particles is topological i.e. it is not a sum of four momenta. Indeed, in order to find the resultant cluster $\eta$ and $\phi$ after the merging of two clusters, the topological algorithm can only compute the average of the particles $\eta$ and $\phi$ belonging to the two clusters. Moreover, if $R_{ij} > 1$ for each particle $i, j$, it is not possible to find the real minimum between all the $d_i$. Therefore, the choice of minimum $d_i$ and thus the choice of the jet $i$ to put in the list, is not fully motivated.

![Figure 4.6: Block diagram for the topological $k_t$ jet algorithm implementation developed in this work.](image)
4.3.3 Jet parameters

In this section the set of parameters chosen to describe a jet, which are assigned by the
topological algorithms once the jets are found, is described. In order to better explain these
parameters, some examples and typical distributions obtained from simulation studies at
particle level, are reported. Regardless of the jet algorithm used for jets reconstruction,
jets are characterized by four parameters to be properly used according to the analysis

- \((\eta, \phi)\) position of the jet centroid.
- Number of charged particles forming the jet (jet particle/track multiplicity).
- A shape parameter describing how much the particles forming the jet are close to
  the jet center.

The \(\eta\) and \(\phi\) distributions of the reconstructed jet centroid, obtained at particle level with
the Di-jet Monte Carlo sample described in section 4.2, are shown in fig. 4.7. Looking at
the top pictures of fig. 4.7, it should be noted a small contribution of jet reconstruction
even in the T1 acceptance. Indeed, even if the partons are generated with pseudorapidity
\(5.0 < \eta < 6.5\), the hadronization processes can also interest pseudorapidities regions
beyond this range.

The shape and the particle/track multiplicity parameters are quantities that can be
used to perform appropriate jet selection. In order to describe the shape parameter, the
function \(S(r)\) is introduced (see fig. 4.8). This function is defined as the number of the
jet particles with distance less than \(r\) from the jet centroid, divided by the number of
jet particles with distance less than \(R\) from the jet centroid. In particular the jet-shape
parameter is defined as the value of \(S(r)\) evaluated at \(\rho = r\) where \(\rho\) and \(R\) are two
quantities to be set according on the events studied: \(R\) is the mean radius dimension of
the jet cone, \(\rho\) is a number less than \(R\). Fig. 4.9 shows the fraction of charged particles at
distance less than \(r\) from the leading (top) and secondary (bottom) jet’s center obtained
with cone (left) and \(k_t\) (right) topological algorithm. Fig. 4.9 is obtained by using all the
events with reconstructed jets in the Monte Carlo sample; it shows that about 70% - 80% of
charged particles fall inside a cone of radius 0.5. Therefore, a reasonable value of \(\rho\), which
4.3 Topological jet algorithms

Figure 4.7: $\eta$ (top) and $\phi$ (bottom) jet center distribution obtained with topological cone (left) and $k_t$ (right) algorithm. Only charged particles in T1/T2 acceptance are selected.

Figure 4.8: Graphical representation for the charged jet-shape $S(r)$: this function is given by the number of jet particles at distance less than $r$ from the jet axis, normalized to the total number of particles of jet particles with distance less than $R$ from the jet axis. The jet shape parameter, used to characterize a jet, is given by the value of this function evaluated at $\rho = r = 0.5$. 
is used in this work for Di-jet events studies, is $\rho = 0.5$. The values of leading jet shape reconstructed by both the algorithms is shown in fig. 4.10 (bottom). The distribution of jet charged particle multiplicity is shown in fig. 4.10 (top) in which, in order to reject a big fraction of fake jets, only jets made by at least two particles are shown. For completeness, $R$ is set here to 0.7 but motivations for this choice will be explained in section 4.4.1, where also the choices of all the other jet algorithm working parameters presented in the previous sections will be motivated.
4.3 Topological jet algorithms

Figure 4.9: The function \(S(r)\) i.e. the fraction of charged particles falling at distance \(r\) from leading (top) and secondary (bottom) jet axis with respect to the total number of particles falling within a radius \(R\) from the jet axis; results obtained with cone (left) and \(k_t\) (right) algorithm.

Figure 4.10: Jet particle multiplicity (top) and leading jet shape (bottom) obtained with cone algorithm (left) and \(k_t\) algorithm (right). Only jets with at least two particles are shown.
4.4 Algorithm setting

In this section the algorithm parameters for both topological cone algorithm and \( k_t \) algorithm are assigned and justified on the ground of simulation results obtained at particle level (see subsection 4.4.1). Moreover, a system of cuts on the jet particle/track multiplicity and the jet shape, which will be used in all the following sections, is established (see subsection 4.4.2).

4.4.1 Algorithm parameters

In order to assign reasonable values to the algorithm parameters, preliminary studies at particle level have been performed. Fig. 4.11 shows the distance \( \Delta R = \sqrt{(\eta_q - \eta_P)^2 + (\phi_q - \phi_P)^2} \) between a charged particle and its nearest parton in the \( \eta - \phi \) plane, where the index \( P \) is used for particles and \( q \) for quarks. Most of the particles are within a distance of \( \Delta R \sim 1 \); this information is used to set to 0.7 the search radius \( R \) of the cone jet algorithm and to 1 the \( D \) parameter used in the \( k_t \) algorithm (this is also usually the choice used in literature, for traditional jet algorithms [48]). However, more detailed studies about the dependence...
of jet reconstruction efficiency on the internal algorithm parameters have been done and are here presented. For instance, fig. 4.12 shows the dependence of the mean number of jets reconstructed by the $k_t$ algorithm when varying the parameter D. It is clear that the $k_t$

algorithm jet reconstruction efficiency increases if the value of $D$ becomes larger. Indeed, increasing $D$, more particles can be clusterized in a single jet and it is more probable for the reconstructed jets to pass the cut on the jet particle/track multiplicity\(^1\). The mean number of reconstructed jet does not increase with $D$ indefinitely: if $D$ is chosen too big, the algorithm will clusterize most of the particles in a single jet. Fig. 4.13 shows the mean distance between the leading (left) and secondary (right) jet from the nearest parton, which gives an idea of the “resolution” by which the jet is reconstructed. From that, it is also evident that the mean distance parton - jet centre (for both leading and secondary jet as well) does not change too much by varying $D$, hence the choice of $D$ is performed mainly on the base of fig. 4.12. As already reported, the value of $D$ adopted hereafter is 1.0. Similar studies have also been done for the cone algorithm. Fig. 4.14 (obtained with a cell size of 0.4) shows the Di-jet efficiency as a function of the cone algorithm searching radius $R$. The Di-jet efficiency is obtained as the number of events with at least two re-

\(^1\)As stated in next subsection, only jet containing at least two particles/tracks are included in the final jet list.
Figure 4.13: Mean distances of leading (left) and secondary (right) jet centre from nearest parton as a function of D.

Figure 4.14: Mean number of Di-jets reconstructed per event with the cone algorithm, as a function of R.

constructed jet normalized to the total number of events generated. Fig. 4.14 shows that the Di-jet efficiency has little improvement when increasing $R$. The value 0.7 is adopted
as reference for the next studies\textsuperscript{2}. Another important parameter for the cone algorithm is the cell size: as shown in the left picture of fig. 4.15, the cone algorithm doesn’t show

![Graph showing jet reconstruction efficiency varying cell size](image1)

**Figure 4.15:** Left: Two jet reconstruction efficiency varying the cell size. Right: Stability of the cone algorithm shifting the grid along the cell diagonal (at fixed cell size).

strong dependency of Di-jet reconstruction efficiency on the cell size so a value of 0.4 is used. Another check on the cone algorithm stability has been performed by studying the Di-jet reconstruction efficiency obtained by shifting the cells grid in the direction of the cell diagonal: fig. 4.15 (right) shows the stability of the algorithm, confirming a reduction of “collinear sensitivity” thanks to the application of the MidPoint procedure.

### 4.4.2 Jet selection cuts

In order to understand the algorithm performance presented in sections 4.5.1 and 4.5.2, it is necessary to remark the default jet selection cuts implemented in the following studies, which are the same at particle and track level:

- only jets with at least two particles/tracks are included in the final jet list;

- a shape cut (where a jet is kept only if its jet shape is above a prefixed threshold)

\textsuperscript{2} The most appropriate $R$ value, as well as the $D$ value for the $k_t$ algorithm, for a given analysis is expected to be derived by proper studies related to it.
4.5 Results and comparisons

can also be required. However, working only with charged particles/reconstructed tracks, there are lots of Di-jet events (about 30%) with less than 4 charged particles in the T2 region, as shown in fig. 4.16.

As a consequence of that, a great loss of clear Di-jet events, “naked eye recognized” by looking at the particle $\eta - \phi$ distribution of generated event, are found if a shape cut is introduced; this feature is summarized in fig 4.17 which shows the mean number of jet reconstructed per event without the shape cut (left) and requiring a jet shape bigger than 0.5 (right): this picture shows that the Di-jet reconstruction efficiency, defined as the fraction of events in which at least two jets are reconstructed, get worse of about 5%-10% if the shape cut is introduced.

![Number of Charged Particles in T2 (di-jet) and Number of Charged Particles in T2 (SD)](image)

Figure 4.16: Number of charged particles for Di-jet (left) and Single Diffractive (right) events within the T2 acceptance.

In future works, more sophisticated shape cuts should be studied using combined information of jet multiplicity and jet shape. Hereafter, if not otherwise stated, this type of shape cuts are not applied.

4.5 Results and comparisons

The performance of both topological jet algorithms described in the previous chapters are presented in this section. These studies are performed at particle level (subsection 4.5.1) and at track level (subsection 4.5.2) both on Di-jet and Single Diffractive events.
4.5 Results and comparisons

4.5.1 Performance at particle level

In this subsection some features of the two topological algorithms are presented with studies at particle level by using the default parameter setting described in subsection 4.4.2 and 4.4.1. This work is not focused on a particular type of analysis, hence the aim of this section is to present the topological jet algorithms as potential tools to be optimized on the particular analysis needs. The number of jet reconstructed per Di-jet event is shown in fig. 4.18 for cone algorithm (left) and $k_t$ algorithm (right). From that, the Di-jet reconstruction efficiency shown is around 35 - 40% and for the cone algorithm it seems to be slightly more efficient. This difference can partially be explained looking at the distribution of the particles found in some particular events, as reported for example in fig. 4.19. The left plot is done by using the cone algorithm without any shape request; here we can see that every pair of particles is grouped thanks to the MidPoint procedure which starts searching cone in the middle of the particle pairs; the resulting cone, having two particles inside, passes the multiplicity cut. The plot in the right is the result obtained using $k_t$ jet...
4.5 Results and comparisons

Figure 4.18: Mean number of reconstructed jets per event in T1-T2 (no shape cut) is shown for the cone algorithm (left) and the $k_t$ algorithm (right).

Figure 4.19: Selected Di-jet events reconstructed by cone algorithm (left) and missed by $k_t$ algorithm (right).

algorithm on the same event which shows that a pair of particles, having distance bigger than 1, are not clusterized. This effect can of course be reduced by requiring a bigger value of the $D$ parameter when the $k_t$ algorithm is used. However, as expected, the main
problem which limits reconstruction efficiency is an algorithm-independent effect: it is the difference between charged jets and real jets due to the fact that neutral particles are not detected. An example that shows this effect, is reported in fig. 4.20, representing two particular Di-jet events reconstructed by using the cone algorithm with (right) or without (left) the addition of neutral particles. On top-left of this figure are shown two “close” (roughly with distance less than two) jets where all the charged particle are falling in the the same half plane, loosing the characteristic topology expected for a Di-jet event. Moreover, as shown on the bottom-left of fig. 4.20, it is possible to have a configuration of charged particles too disperse, which results in the impossibility of particles grouping. It is interesting to note that these problem is often resolved when the algorithm is applied also adding the neutral particles (right pictures in fig. 4.20). Other examples of Di-jet
4.5 Results and comparisons

Events successfully reconstructed by adding the neutral particles in the event are shown in fig. 4.21.

![Di-jet event reconstruction](image)

Figure 4.21: Two particular events which could be correctly reconstructed only with the addition of neutral particles. The jet reconstruction is obtained with the cone algorithm by using only charged particles (left) and including neutral particles (right).

Other expected issues are found whenever a cut on particles $\eta$ is requested as shown in fig. 4.22, where the reconstructed jets for two particular events are obtained without any selection on particle $\eta$. The figure confirms an expected geometrical issue, due to the limited acceptance of TOTEM inelastic detectors and the possibility for the particles to fragment outside T2. Because of the uncovered rapidity gap between T1 and T2, in some events (see fig. 4.22, right) two jet reconstruction is not achieved. However, the use of both the inelastic detectors is important (see fig. 4.22, left) in order to obtain a better reconstruction efficiency even if the partons are generated outside the T1 acceptance. Due
to the fact that many events completely fail to reproduce two distinct groups of particles, it is possible to explain a features of the fig. 4.23 where the azimuthal separation between the leading and the secondary reconstructed jet center is plotted. As expected, fig. 4.23 shows a peak around $\pi$ but contributions from lower values of azimuthal separation (the reconstruction at $\Delta \phi > \pi$ represent an effective jet separation of $2\pi - \Delta \phi$) are also present, due to the kind of close events already shown in fig. 4.20. However, as fig. 4.24 shows, the

**Figure 4.23:** Azimuthal separation between leading and secondary jets as obtained with cone (left) and $k_t$ (right) algorithm.
4.5 Results and comparisons

algorithms allows to obtain quite good resolutions, i.e. the distance of the reconstructed jet center from nearest parton is, on average, quite small. Using this algorithm parameter setting, a better result is obtained with the $k_t$ algorithm. Distances bigger than 1.2 present in fig. 4.24 are due to “bad shaped-events” as said before and as put in fig. 4.25. Since at least two particles per jet are required, it is important to remark that the algorithm has a chance to find 2 jet only for Di-jet event with 4 or more particles in the T1/T2 acceptance. Consequently, many reconstructed events are expected with no jet or only one jet, as shown, for instance, in 4.17. The correct implementation of the algorithm is checked by performing the same studies shown in fig. 4.17 but now requiring at least 4 particles in the events: the results are shown in fig. 4.26 where it can be noted that Di-jet efficiency gets better of about 15%, with respect to the left pictures in fig.4.17. Thankfully, both the algorithms are quite able to discern real two jet events from SD events as shown in

![Figure 4.24: Distance between leading (left) and secondary (right) jet centre and nearest parton as obtained with cone (top) and $k_t$ (bottom) algorithm.](image-url)
4.5 Results and comparisons

Figure 4.25: Example of event with a jet centre reconstructed at more than 1.2 from nearest parton.

Figure 4.26: Number of jets reconstructed in T1-T2 per event. Only event with at least 4 charged particle in T1-T2 are considered. Left: cone algorithm results; right: $k_t$ algorithm results.

fig. 4.27. In this figure the jet reconstruction efficiency is obtained applying the algorithm on SD events: in the ideal case, the algorithms should give no jet reconstruction in all the events instead our algorithms reconstruct Di-jet events with an efficiency around 15-20%. However, it is important to remark that big improvements in background rejection could be achieved, optimizing the cuts and the algorithm parameters on the particular event studied.
4.5 Results and comparisons

Figure 4.27: Results for cone algorithm (left) and $k_t$ algorithm (right) when applied on Single Diffractive events. The distribution of the average number of reconstructed jets per event is shown.

4.5.2 Performance at track level

In this subsection the performance of the topological jet algorithms are studied at charged track level. 900 Di-jet event and 640 Single Diffractive events are generated, with the usual minimal cuts ($5.0 < \eta < 6.5; p_t > 3$ GeV/c for the outgoing partons) including the simulation of T1 and T2 detectors. The same studies shown in section 4.5.1 are repeated although this time the algorithms will clusterize the T1 and T2 reconstructed tracks. As already reported in 3.5.2, in order to use jet algorithms on the reconstructed tracks, a preliminary track selection becomes mandatory. For the algorithms it is fundamental to work only with primary tracks in order to preserve some relation between the charged particle density and the particles energy in the $\eta - \phi$ plane. In the studies presented here, the standard selection cuts system already described in 3.5.2 will be used. It is expected that if the $\sigma_{\Delta Z}(\eta)$ curve (see fig. 3.11, bottom) obtained for 50 GeV pions are used for the cut on track $Z_0$, the charged hadrons having energy much smaller than this value undergo a bigger rejection with respect to the higher energy component. In a certain sense, choosing the 50 GeV curve (instead of the 10 GeV curve) allows a more likely selection of the high energy particles, as a traditional jet algorithm should do. In order to make
The user can optimize, according to his analysis needs, the track selection criteria as well as all the internal algorithm parameters and the thresholds for jet selection cuts. Both the $\sigma_{\Delta Z}(\eta)$ curves obtained with pions at $E = 10$ GeV and $E = 50$ GeV presented in fig. 3.11 are available and the $Z_0$ cut can be set at the desired constraining level by changing the number $n$ which appears in the $Z_0$ cut condition:

$$Z_{0\text{Tr}} < n \sigma_{\Delta Z}(\eta_{\text{Tr}})$$  \hfill (4.2)

where $Z_{0\text{Tr}}$ and $\eta_{\text{Tr}}$ are respectively the reconstructed track $Z_0$ and $\eta$ (see section 3.5.2).

All the algorithm handles which can be set by the user are summarized in table 4.1.

<table>
<thead>
<tr>
<th>Jet Finder</th>
<th>Handle</th>
<th>Present value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cone algorithm</td>
<td>Grid minimum $\eta$</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Grid maximum $\eta$</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>Grid minimum $\phi$</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Grid maximum $\phi$</td>
<td>$2\pi$</td>
</tr>
<tr>
<td></td>
<td>Cell size $\Delta \phi$</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Cell size $\Delta \eta$</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Search Radius ($R$)</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td>$f_{\text{merge}}$</td>
<td>0.5</td>
</tr>
<tr>
<td>$k_t$ algorithm</td>
<td>$D$ value</td>
<td>1.0</td>
</tr>
<tr>
<td>Both (Track selection)</td>
<td>T2 minimum $\eta$</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>T2 maximum $\eta$</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>T1 minimum $\eta$</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td>T1 maximum $\eta$</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{\Delta Z}(\eta)$ curve energy</td>
<td>50 GeV</td>
</tr>
<tr>
<td></td>
<td>$n$ for $Z_0$ cut (see eq.4.2)</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>$\chi^2$ cut</td>
<td>1.0</td>
</tr>
<tr>
<td>Both (Jet selection)</td>
<td>Shape Cut</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>Shape Radius</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Track Multiplicity Cut</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.1: Jet algorithms handles available for jet reconstruction at track level.

The distribution of the $\eta$ (top) and $\phi$ (bottom) jet centers, reconstructed with tracks in T1-T2 is shown in fig. 4.28, to be compared with the correspondent fig. 4.7 obtained at particle level. Even with the use of strict selection cuts on the tracks, a little degradation
4.5 Results and comparisons

Figure 4.28: \( \eta \) (top) and \( \phi \) (bottom) jet center distribution, obtained at track level with the cone algorithm (left) and the \( k_t \) algorithm (right).

on the algorithms performance are expected. The main issues which appear when going from charged particle to T1/T2 track level are due to magnetic field, secondary particles, and inefficiencies in track reconstruction or selection. More precisely:

- the magnetic field deviates primary low energy particles, changing mainly their original \( \phi \) coordinate. 10 GeV pions suffer an azimuthal shift of about 20° and a pseudorapidity shift of about 0.05 (see subsection 3.5.3). The inclusion in the event of these tracks fakes the topology of the event and can lead to jet reconstruction in \( \eta \)-\( \phi \) regions which are distant from the correspondent \( \eta \)-\( \phi \) parton coordinates;

- tracks coming from secondary particles which are not rejected by the cuts currently used, are treated on the same footing as the primary tracks. Therefore it is im-
important to have a very strict primary track selection because if a region includes secondary tracks, the position and the multiplicity of these fake primary tracks are not representative of the real energy flow in that region;

- inefficiency on primary track reconstruction or selection reduces the primary track density; as a consequence, fewer jets are reconstructed.

The effect of the magnetic field and of the inclusion of secondary tracks is expected to affect the distribution of azimuthal separation between leading and secondary jets and the distributions of the distance between the reconstructed jet and the nearest parton. The distribution of the azimuthal separation is reported in fig. 4.29 which should be compared to fig. 4.23. The distance between the jets and the nearest quark are instead shown in picture 4.30 which has to be compared with fig. 4.24. As shown in fig. 4.31 (top), the Di-jet event reconstruction efficiency (fraction of events with at least two reconstructed jets) obtained at track level is about 25%-30%. This value, with respect to the efficiency obtained at particle level (see fig. 4.18), is about 10% lower. However, it should be noted that even the rejection of background events, as for example Single Diffractive events is more efficient (see fig. 4.31 (bottom) and 4.27). The fake Di-jet rates goes from 15%-20% obtained at particle level to less than 5% at track level. Even if the algorithms are not yet optimized in order to obtain the best efficiency for Di-jet event and the best rejection for Single Diffractive events, going from charged particle level to track level, the better
rejection of fake jet reconstruction in Single Diffractive event is more evident than the efficiency loss in Di-jet events. This behaviour is explained as consequence of the fact that the primary track reconstruction efficiency is worse for low energy particle and the mean particle energy for Single Diffractive event is smaller than for Di-jet events, as shown in fig. 3.13. The observed performance of the current jet algorithms run at track level, as shown in these preliminary studies, gives an idea of their big potentiality when used in analyses involving jets in the very forward region. These algorithms are so expected to provide an important tool to these analyses.

Figure 4.30: Distances between leading (top) and secondary (bottom) jet center and nearest parton as obtained with cone (left) and $k_t$ (right) algorithm.
Figure 4.31: Mean number of jet reconstructed in Di-jet (top) and Single Diffractive (bottom) events, as obtained with cone (left) and $k_t$ (right) algorithm.
4.5 Results and comparisons
Conclusions

The TOTEM Experiment at the CERN LHC is designed and optimized to measure the total $pp$ cross section, to study the nuclear elastic $pp$ cross section and to perform a comprehensive physics programme on diffractive dissociation processes partially in cooperation with the CMS Experiment. TOTEM will be ready for data taking at the very beginning of the LHC start.

The measurement of the total $pp$ cross section requires the simultaneous detection of the elastic scattering at very low squared four-momentum transfer (down to $-t \sim 10^{-3}$ GeV$^2$) and the evaluation of the total inelastic rate. The measurement of the inelastic rate is performed by two tracking telescopes, T1 and T2 (based on gas detectors), which detect charged particles scattered at $18 \text{ mrad} < \theta < 90 \text{ mrad}$ and $3 \text{ mrad} < \theta < 10 \text{ mrad}$ respectively. They also provide the reconstruction of primary vertex, so that beam-beam events can be disentangled from background. In this thesis, a track reconstruction algorithm for the T2 detector, which will be used for inelastic rate measurement and vertex reconstruction, has been developed. In the first part of the thesis, the algorithm which allows the reconstruction of the geometrical position of the ionizing particle inside the detector starting from the electrical signal generated in the detector read-out board (hit reconstruction), is presented. A tracking algorithm, which allows to reconstruct the charged particle trajectory through the detector, has been then developed and its performance tested. As T2 has to provide the reconstruction of the primary vertex of the event, studies about rejection of secondary tracks (generated by interactions of primary particles with material in front of or around T2) have also been performed. In order to select only primary tracks, a set of cuts to be applied on the reconstructed track parameters is proposed. In particular the effect of the beam pipe on particle multiple scattering and secondary particle production has been
studied into details. This led to the observation of a region in the T2 acceptance, around $|\eta| \sim 5.53$, with very low possibility of primary particles detection. Preliminary studies have shown a track reconstruction pseudo-efficiency (mean number of reconstructed tracks for primary charged pion) of about 100% without primary track selection criteria. When such selection cuts are applied, the primary track pseudo-efficiency is well above 80% everywhere, excluding the acceptance region around $|\eta| \sim 5.53$. However, especially at high $\eta$, the primary track reconstruction pseudo-efficiency can surely be enhanced by using less strict cuts on the reconstructed track parameters. Optimization studies in such direction are not yet been performed. It has been also shown that photons produced in neutral pion decays, for pion pseudorapidity $|\eta| > 4.5$, can interact with the beam pipe or with the material around the beam pipe, producing $e^+ - e^-$ pairs which sensibly increase the number of secondary tracks reconstructed in T2.

The primary track selection capability is also tested in the second part of the thesis, where the possibility of jet reconstruction by using charged particle tracks detected with the TOTEM inelastic telescopes is investigated. Jet reconstruction in the forward region can provide TOTEM with the capability of studying many other interesting physics processes in addition to its standard measurements. Two innovative algorithms (cone-like and $k_t$-like) have been developed in order to reconstruct jets by using only the information on charged particle density in the $\eta$-$\phi$ plane. The algorithms have been developed and tested by using Pythia Di-jet events (with quarks generated at $5.0 < \eta < 6.5$ and with $p_t > 3$ GeV/c) and Single Diffractive events (used in order to estimate some background effect). The two algorithms, developed at the charged particle level and tested at the reconstructed track level, show a similar performance. The results presented in this thesis are obtained by using only charged particles falling in the T1/T2 region even if the algorithms can in principle be used in any desired $\eta$-$\phi$ region. Since the jet algorithms implemented at track level use only primary tracks, the comparison of the jet algorithm performance obtained at particle and track level also allows to test the primary track selection criteria in a very high track multiplicity situation. The output of the algorithm is a list of jets, where each jet is characterized by the following informations: $\eta$, $\phi$ of the center, number of charged particles and jet shape. The jet algorithm parameter setting, as well as the jet selection cuts, can be chosen and optimized according to the analysis needs. In this thesis, some
preliminary studies on the dependence of the algorithm performance from the internal parameters and from the cut setting on the reconstructed jets are reported. Even without an optimal setting of the algorithm parameters (as the work reported here is not focused on a particular analysis), it was found a Di-jet event reconstruction efficiency of 35-40% and 25-30% at particle and track level respectively; while the fake Di-jet event reconstructions efficiency obtained in Single Diffractive events is 15-20% and below 5% at particle and track level respectively. These preliminary studies give an idea of the big potentiality for these jet algorithms to provide an important tool to analyses involving jets in the very forward region. Furthermore, for an optimal jet reconstruction, these algorithms can in principle be combined with the energy information from the CMS CASTOR calorimeter.
Acknowledgment

I would like to express my sincere appreciation and gratitude to Prof. Giuseppe Latino for his tireless assistance, support and guide during the thesis work. Thanks for providing me with an optimum working environment and for invaluable suggestions and teaching. Thanks also for all the hours he has spent in correcting my poor English: I’m indebted to him.

I’m also pleased to thanks Prof. Angelo Scribano and Prof. Stefano Lami for allowing me to work in the TOTEM experiment. It has been a great experience which allowed me to learn a lot of things and where I found a very stimulating and clever group. I would like also to thank the CERN colleagues, in particular Dr. F. Ferro, Dr. J. Kaspar and Dr. H. Niewiadomsky for their help and assistance in several technical problems. Special thanks also to Dr. G. Bonnoli and Dr. E. Oliveri for their many sincere advices and useful hints.

At the end of these years of studies, I want also to thank the university friends, mainly T. Squartini, for sharing with me many hours of lessons; and my professors, especially M.A. Ciocci, E. d’Emilio, E. Mariotti, V. Millucci, R. Paoletti and N. Turini for their teaching and attention.

On the non-scientific side, I would like to express my deepest gratitude to Elena, constant benchmarks during these years of studies, for her understanding and love.

I dedicate this thesis to my mother Lucia and to the memory of my father Roberto.

Your constant encouragement and love will never be forgotten.
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).


[37] Erik Brucken, *private communications*.

[38] Eraldo Oliveri, *private communications*.


[40] Fabrizio Ferro, *private communications*.


