Universitá degli Studi di Torino
Scuola di Dottorato in Scienza e Alta Tecnologia

Tesi di Dottorato di Ricerca in Scienza e Alta Tecnologia.
Indirizzo: Fisica e Astrofisica

THE CMS SILICON TRACKER: FROM THE PERFORMANCE STUDY IN COSMIC RUNS TO THE MEASUREMENT OF THE MOMENTUM RESOLUTION FROM THE Z LINE-SHAPE IN EARLY DATA

Maria Assunta Borgia

Tutor: Prof. Marco Costa

XXII Ciclo, Gennaio 2010

Contents

Introduction

The Large Hadron Collider (LHC) is finally entering the collision mode after about 20 years of R&D, design, construction and commissioning. For a safe commissioning of the machine, the beam energy will be rump up in several steps: after few "pilot" runs at 450 GeV and 1.18 TeV per beam in 2009, the LHC is expected to rump up the beam energy to 3.5 TeV by February 2010 and to run with these energies until a significant data sample has been collected and experience in running the machine has been gained. At the end of 2010, the LHC will circulate lead ions beams for the first time. In 2011 the LHC will shut down and work will begin to bring the machine towards 7 TeV per beam. Four detectors, CMS (Compact Muon Solenoid), ATLAS (A ToroidaL ApparatuS), LHCb (The Large Hadron Collider beauty experiment) and ALICE (A Large Ion Collider Experiment) are ready for the data-taking phase with intense detector commissioning works underway.

The much awaited discoveries at the LHC have to be preceeded by the rediscovery of the Standard Model (SM) physics, and early discoveries, though possible, will be challenging. On the other hand, from early phase on, i.e., starting from an integrated luminosity of mere 1 pb^{-1}, various SM processes can be studied; in particular, heavy quarkonia states (e.g. J/Ψ and Υ) will shine well above background in the low mass range of the invariant mass distribution of dilepton final states.

Electroweak processes of W and Z productions, having reasonably large cross sections and clean leptonic decay mode signatures (e/μ), are part of the early physics studies at the LHC and they will be studied as well at all luminosities. These will be also used as standard candles for a large variety of LHC measurements.

Besides being fundamental backgrounds to new physics searches and providing a raw luminosity monitoring, the study of Z and W events will improve the knowledge of the Parton Distribution Functions (PDFs) in an unexplored kinematic region. These analyses hold the key for all the subsequent physics searches because they provide the way to control the main experimental and theoretical systematics at the LHC.

In addition, before any discovery can be claimed, a well-founded control of the detector and reconstruction performances is needed: the muon momentum scale and resolution must be measured from data. Because of the large production cross-section and the detailed knowledge of its mass resonance curve from LEP experiments, the Z boson decay into a muon pair provides a perfect process to perform both the scale calibration and the resolution measurement. The Z boson events analysis provides also a fast way to tune the Monte Carlo simulation, in order to obtain from the Monte Carlo the most realistic description of real data.

This thesis is focused on prospects for the measurement of the muon momentum resolution from data and the tuning of the Monte Carlo simulation with the first 10 pb^{-1} integrated luminosity of $\mathrm{pp} \rightarrow \mathrm{Z} + \mathrm{X} \rightarrow \mu\mu + \mathrm{X}$ events with CMS detector.

Studies have been performed on Monte Carlo samples generated with `SHERPA` including a realistic detector simulation and addressing the most relevant systematic effects, focusing on the CMS Silicon Strip Tracker, as I dedicated a large part of my Ph.D. to the calibration of this very complex object.

After a general introduction in Chapter 1 about the Electroweak physics at the LHC and the rediscovery of the Standard Model in CMS, from Chapter 2 onwards the thesis describes my Ph.D. work.

In Chapter 2, a description of the CMS Silicon Strip Tracker is given, starting from the layout and going ahead with the description of the calibration and reconstruction algorithms.

Since few years the CMS detector is being commissioned using cosmic muons. These have proved to be very valuable to test hardware operation and software algorithms. I analyzed the cosmic ray data collected in the last three years for the Tracker commissioning to perform calibration and to study local reconstruction performances. The results obtained are shown in Chapter 3 and have been published in [1], [**?**], [2].

The measurement of the transverse momentum of charged particles has a vital importance at CMS. In particular the transverse momentum and its resolution are the most subjected to the details of the detector integration and operation, being highly sensitive to the precise alignment of the silicon tracker and of the muon chambers, and in general to their performance. The silicon tracker has a leading role in the momentum resolution of muons up to 200 GeV: its calibration and alignment conditions will be a pivotal requirement during the real data taking in order to have the best possible resolution. An algorithm has been developed to measure the muon momentum scale and the resolution exploiting well known di-muon resonances (Z, Υ, J/ψ) on a wide mass range. I used this algorithm on a Monte Carlo sample of $\mathrm{Z} \rightarrow \mu\mu$, comparing the results obtained on the sample reconstructed in ideal and realistic detector condition. In the realistic scenario I introduced the residual miscalibrations and misalignments which are expected to still affect the data after the collection of an integrated luminosity of 10 pb^{-1}. A total of about 5000 $\mathrm{Z}\rightarrow \mu\mu$ events are expected to be collected in this first phase, and in Chapter 4 I will show how it will be possible to measure the resolution directly from these data.

Finally, for precision measurement and for early discoveries as well, it is very important that the Monte Carlo simulation is able to reproduce the data with good approximation. In the last part of Chapter 4 a method is thus shown to tune the Monte Carlo simulation of the CMS detector with early data in order to have the most realistic description of the detector and to obtain from the Monte Carlo the most precise description of data. This is a fundamental requirement to be enabled to perform new physics searches.

Chapter 1

Standard Model physics at the Large Hadron Collider (LHC)

1.1 Motivations for LHC

The Standard Model (SM) of electroweak interaction is the theory that nowadays describes best the elementary particle physics, since the observation of the neutral current interaction in 1973 with the Gargamelle detector [3] and the observation of the W and Z bosons in 1983 at the UA1 experiment [4] gave the experimental proof of it. Since then, the LEP and SLC experiments gave confirmations of the Standard Model predictions from 1989 to 2000, measuring with better than per-mill precision the W and Z bosons properties: their masses, their widths, their couplings with fermions and among themselves. Tevatron observation of top quark complemented finally these measurements. The last piece missing of the SM is the Higgs boson, which should be the remnant of the scalar field that provides mass to the particles [5]. Precision measurements of the electroweak observables (m_W, α_{em}, θ_W) indicate a light Higgs boson. The accuracy reached in the measurement of these observables in fact requires that, if one relates them among each others, genuine electroweak corrections Δr should be included, namely:

$$m_W^2 = \frac{\pi\alpha_{em}}{G_F\sqrt{2}}\frac{1}{sin^2\theta_W(1-\Delta r)} \tag{1.1}$$

where the quantum corrections have a quadratic dependence on the mass of the top quark m_{top} and a logarithmic dependence on the Higgs mass: $\Delta r = f(m_{top}^2, \ln m_H)$. Once m_W, m_{top} and $\sin^2\theta_W$ are measured, m_H can be extracted from a global fit of the electroweak observables, as shown in Fig. 1.1.

LEP has established at 95% confidence level [6] the lower limit on the Higgs mass: 114.4 GeV/c^2. On the other hand, an upper limit around 1.2 TeV is derived within SM requiring that the amplitude for the scattering of longitudinally polarized vector bosons $V_LV_L \rightarrow V_LV_L$ does not violate

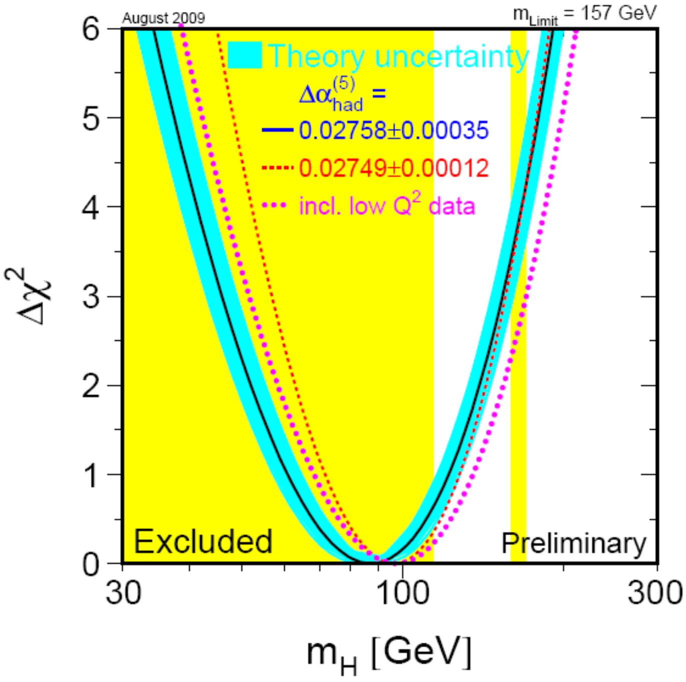


Figure 1.1: *Global fit to electroweak precision measurements done by the LEP Electroweak Working Group for Winter 2008, with the additional Tevatron excluded region.*

unitarity [7]. Moreover, another recent exclusion region at 95% confidence level has been found by the CDF and D0 experiments at Tevatron, in the range of mass 163-166 GeV$/c^2$.

The existence of the Higgs boson is just one of the possibilities to break the symmetry and provide masses to the particles. The Goldstone theorem and the Higgs mechanism do not require the existence of elementary scalars. It is also admitted as a possibility that bound states are responsible for electroweak symmetry breaking. Taking into account that unitarity is nothing but a statement of conservation of total probability, it is evident that it cannot be violated in nature. Violation of perturbative unitarity would imply the SM to become a strongly interacting theory at high energy. If the Higgs mass was large or the Higgs didn't exist, by analogy with low energy QCD, one might expect the presence of resonances in $V_L V_L$ scattering. But unfortunately the mass, spin and number of these resonances are not uniquely determined.

The discovery of the mechanism which gives origin to the masses requires the deep investigation of the energy range from 100 GeV to 1 TeV. For this reason LHC has been designed as a discovery machine for processes with cross sections down to some tens of fb and in the energy range from 100 GeV to 1-2 TeV. This physics goal influenced the main design parameters of the machine:

- being a hadron collider, the colliding entities are the partons, which carry a variable fraction x of the whole hadron momentum. Therefore the center-of-mass energy of the hard scattering process $\sqrt{\hat{s}}$ can span different order of magnitude.

  Considering the center-of-mass energy $\sqrt{s}$ = 14 TeV, and the incoming partons carrying momentum fraction $x_1, x_2 \approx$ 0.15-0.20 of the incoming hadrons momenta, the partonic CM energy is $\sqrt{\hat{s}} = x_1 x_2 \sqrt{s} \approx$ 1-2 TeV, which matches exactly the energy range to be explored.

- the beams are made of protons. This has a fundamental effect, which is the ease with which protons can be accelerated, with respect to electrons and positron (LEP). In fact the energy lost by synchrotron radiation is proportional to γ^4 (where γ = E/m), and this is much smaller for protons than for electrons and positrons. Moreover, the choice to accelerate two protons beams instead of a proton and an antiproton beam permit to obtain high intensity beams (it is easier to accumulate protons with respect to antiprotons) without losing much in the Higgs production, as it is a process dominated by gluon fusion which has almost the same cross section in $p - \overline{p}$ and $p - p$ collisions.

Another main parameter strongly influenced by the physics goal is the luminosity. In order to compensate for the low cross section of the interesting processes, LHC must provide a very high number of collisions. This is achieved with a very short bunch crossing interval (25 ns, which means a frequency of 40 MHz) and a high number of bunches accelerated by the machine (2808 per beam). This will allow to reach a peak luminosity of 10^{34} cm^{-2}s^{-1}. Only about 80% of the bunches will be filled. All this arrangement will permit to accumulate, during the first three years of operation at the design center of mass energy, an integrated luminosity of $\int \mathcal{L}$dt = 20 fb^{-1} per year with $\mathcal{L}$ = 2×10^{33} cm$^{-2}s^{-1}$, and later $\int \mathcal{L}$dt = 100 fb^{-1} per year with $\mathcal{L}$ = 2×10^{34} cm$^{-2}s^{-1}$.

The whole set of parameters of the LHC is summarized in Table 1.1.

The main drawback of this powerful design is that the total event rate will be so high that several interactions overlap in the same bunch crossing, creating the so called *pile up*. It has been estimated that, with a 55 mb cross section for p-p non-diffractive inelastic process predicted by PYTHIA, there will be on average 17.3 events occurring at every bunch crossing, with about 50 charged tracks per interaction. The high bunch crossing frequency, the high event rate and the pile-up of several events in the same bunch crossing dictate strict requirements on the design of the detectors. The bunch crossing of 25 ns and the pile-up of about 20

Parameter	p-p	Pb
CM energy (TeV)	14	1148
Number of particles per bunch	1.1×10^{11}	~8×10^{7}
Number of bunches	2808	608
Design luminosity (cm^{-}2s^{-}1)	10^{34}	2×10^{27}
Luminosity lifetime (h)	10	4.2
Bunch length (m)	53	75
Beam radius at interaction point (μm)	15	15
Time between collisions (ns)	24.95	124.75
Bunch crossing rate (MHz)	40.08	0.008
Circumference (km)	26.659	26.659
Dipole field (T)	8.3	8.3

Table 1.1: *LHC parameters for p-p and Pb-Pb collisions.*

events per crossing influence the time response and the readout electronics, which have to be very fast. The presence of pile-up means instead that a high granularity is necessary to avoid the overlap of particles in the same sensitive elements. High granularity in turn means a large number of channels and therefore high cost. Moreover, LHC detectors will also have to stand extremely high radiations doses, so all the electronics will need to be radiation hard. Finally, additional requirements apply to the online trigger selection, that has to deal with a background rate several orders of magnitude higher than the signal rate.

1.1.1 The LHC physics program

In Fig. 1.2 the cross sections and the production rates of interesting processes at LHC are shown as a function of the center-of-mass energy and of the mass of the produced particles. In Table 1.2 the cross section and the number of events produced for a given process per experiment for low luminosity ($\mathcal{L} = 2\times10^{33}$ cm^{-2}s^{-1}) are reported.

Process	σ	Events/s	Events/y	Other machines
$W \rightarrow e\nu$	20 nb	15	10^{8}	10^{4} LEP/10^{7} Tev.
$Z \rightarrow ee$	2 nb	1.5	10^{7}	10^{7} LEP
$t\bar{t}$	1 nb	0.8	10^{7}	10^{5} Tevatron
$b\bar{b}$	0.8 mb	10^{5}	10^{12}	10^{8} Belle/Babar
$\widetilde{g}\widetilde{g}$ (m=1 TeV)	1 pb	0.001	10^{4}	
H (m=0.8 TeV)	1 pb	0.001	10^{4}	
H (m=0.2 TeV)	20 pb	0.01	10^{5}	

Table 1.2: *Expected cross sections and number of events per second and per year for an experiment at LHC.*

As the Higgs cross section increases steeply with the center-of-mass energy while the total cross section (i.e. the background) remains almost

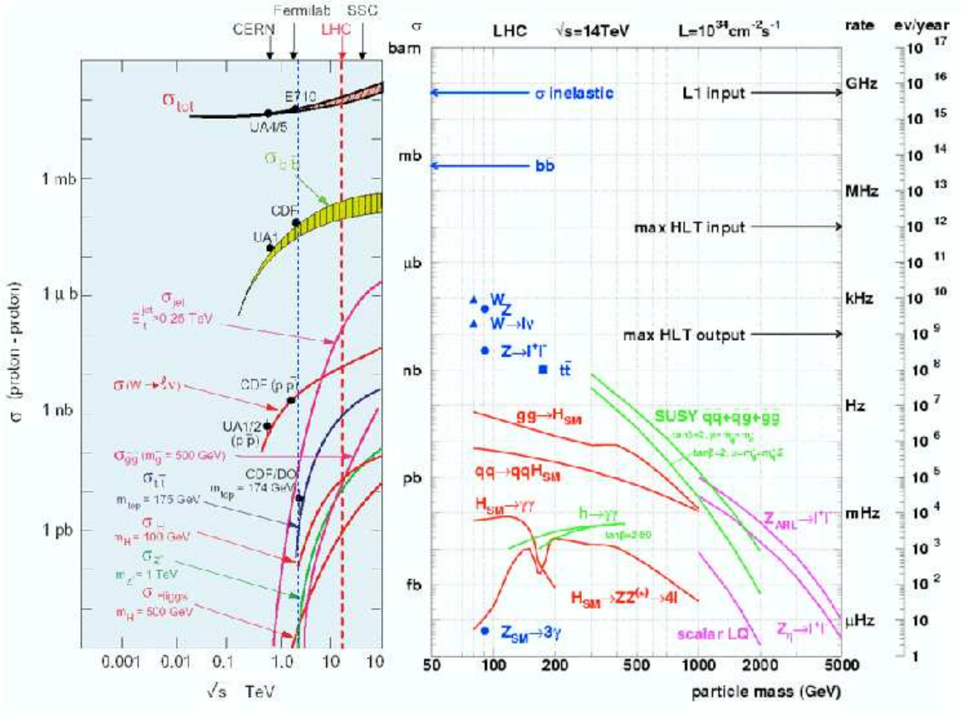


Figure 1.2: *Cross section as a function of the center-of-mass energy (left) and event rates at LHC as a function of the mass of the produced particle (right) for interesting processes.*

constant, the highest center-of-mass energy should be used.

The conception of the Large Hadron Collider had as background idea to reuse the existing 27 km long LEP tunnel to install a new collider, substituting electron beams with proton beams. This idea permitted considerable financial saving, considering that the tunnel itself and many infrastructures (experimental caverns, pre-accelerators) already existed. The drawback is that the maximum center-of-mass energy to 14 TeV is limited by the dimension of the tunnel itself, since the beams must be bent by dipole magnets whose maximum field is currently limited at about 8 T.

One very remarkable aspect of LHC physics is the overwhelming background rate compared to the interesting physics processes: the Higgs production, for example, has a cross section at least ten orders of magnitude smaller than the total inelastic cross section, as shown in Fig. 1.2. This is due to the fact that the bulk of events produced in the proton-proton collisions is either due to the low-$\widehat{p}_T$ scattering, where the protons collide at large distances, or to QCD high-$\widehat{p}_T$ processes of the type:

$$\begin{cases} q_i\overline{q}_i \rightarrow q_k\overline{q}_k \\ q_iq_j \rightarrow q_iq_j \\ q_ig \rightarrow q_ig \\ q_i\overline{q}_i \rightarrow gg \\ gg \rightarrow q_k\overline{q}_k \\ gg \rightarrow gg \end{cases} \tag{1.2}$$

All these events are collectively called “minimum bias”. They constitute the background for other processes in which massive particles are created in the hard scattering, so they are in general considered uninteresting in LHC physics studies. This classification can be in some case misleading, as it includes for example also processes like $b\bar{b}$ which are interesting for B physics.

Another implication of the choice of colliding protons is the span in the energy of the initial state partons, whose energy fraction is given by the Partons Density Functions (PDFs). Two examples of CTEQ4M PDFs for two different Q^2 are shown in Fig. 1.3.

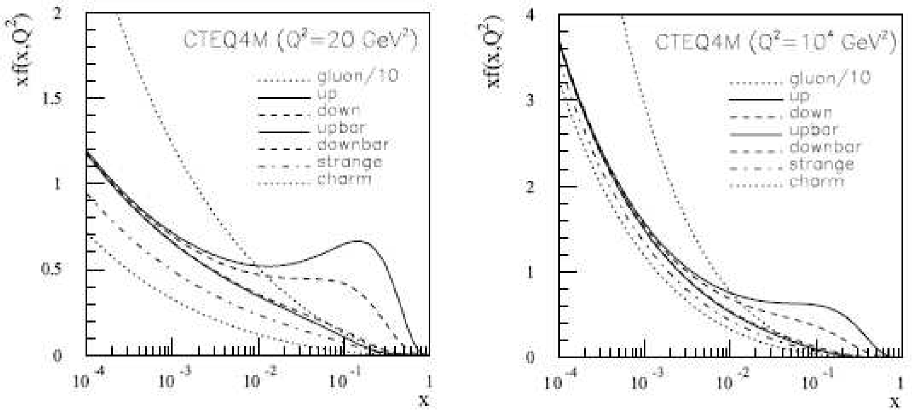


Figure 1.3: *Parton density functions for* Q^2 *= 20 GeV/*c^2 *and* Q^2 *=* 10^4 *GeV/*c^2*.*

The fact that the collisions between the two partons happens at an unknown energy has two fundamental consequences. The first implication is that it is impossible to know which is exactly the total energy of an event, because the proton remnants, that carry a sizable fraction of the proton energy, are scattered at small angles and are predominantly lost in the beam pipe, escaping undetected. Experimentally, it is therefore not possible to define the total and missing energy of the event, but only the total and missing transverse energies (in the plane transverse to the beams). Moreover, the center of mass may be boosted along the beam direction.

This is the reason for which it is necessary to use experimental quantities which are invariant under such boosts as the transverse momentum p_T. Another experimental quantity can be introduced, the *rapidity*, which is defined (choosing the beam direction as z axis) as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \tag{1.3}$$

and it is additive under boosts along the beam axis. Therefore the shape of the dN/dy distribution is invariant under such boosts and this is the reason why it is more convenient to use the rapidity to describe angular distributions.

In the limit of ultra-relativistic particles ($p \gg m$) the rapidity is approximated by the *pseudorapidity*:

$$\eta = -\ln \tan \frac{\theta}{2} \tag{1.4}$$

where θ is the angle between the particle momentum and the z axis. Therefore the pseudorapidity can be simply reconstructed measuring the θ angle, independently on the mass and the momentum measurements. As statistical particle distributions are flat in η for many physics production models (for instance the distribution of charged particles in the underlying event in the central region) it is often useful to define a cone around a given particle in terms of pseudorapidity:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{1.5}$$

This variable is often used in isolation algorithms and in the Monte Carlo truth method.

1.2 Precision Tests of the Standard Model

In addition to the large discovery potential for Higgs bosons and new physics phenomena, like for instance Supersymmetry, experiments at the LHC will also be able to perform precise measurements of many Standard Model processes. The huge cross sections and the high collision energy will allow the exploration of less well tested areas of the Standard Model and to improve significantly on the precision of many parameters. Therefore the detailed experimental study of known Standard Model processes at the LHC will not only prepare the anticipated discoveries but also complement their interpretation.

1.2.1 The structure of the proton

Most of the interesting reactions at the LHC are hard scattering processes of the partons inside the protons. Therefore the accurate knowledge of the parton distribution functions (PDF) of quarks, anti-quarks and gluons is mandatory for the interpretation of the data.

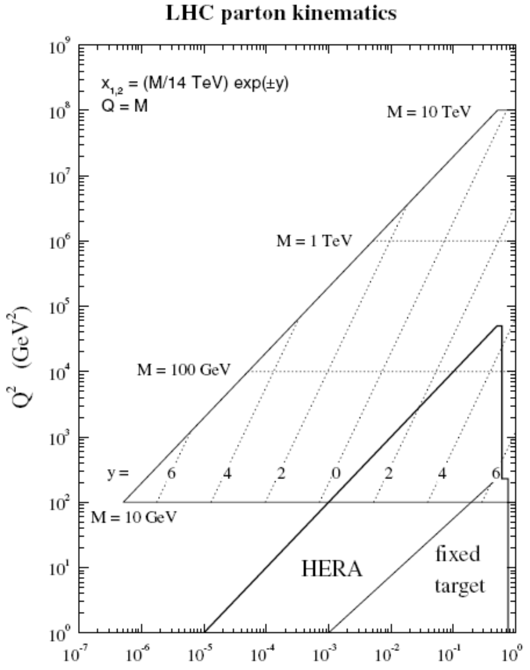


Figure 1.4: *LHC area of interest in the* $x - Q^2$*-plane of the scattering partons compared to the regions accessible at HERA and fixed target experiments.*

The region of the fractional parton momentum x and the momentum transfer Q^2 interesting for LHC physics is shown in Fig. 1.4. This is compared to the area investigated by fixed target experiments and at the electron-proton collider HERA. As the overlap of both regions is small, theoretical models like the DGLAP evolution [8] are required to extrapolate from the experimentally tested regions of low Q^2 to the higher momentum transfers important for the LHC physics. It is therefore desirable to determine the PDFs for all parton species directly from LHC data, to test the extrapolation models and to improve the precision. SM processes provide various means to measure and constrain the PDF through differential distributions like the distributions of transverse momentum p_T and pseudo-rapidity η. Obvious examples are measurements of differential cross sections in W and Z production, but also quark gluon scattering processes like $gq \rightarrow q\gamma$. These reactions have a large cross section and thus lead to small statistical uncertainties already in the initial phase of the experiments. However the achievable precision of such analyses still has to be demonstrated taking into account full detector and background simulations. Because of the high cross section of these processes the analyses are particularly suited for the start-up phase at lower luminosities.

Another issue related to the composite structure of the proton is the so-called underlying event which is understood as the proton-proton interaction without the hard scattering process. The precise modelling of the underlying event is important for the experimental quantities like jet definition and resolution and therefore for the correct interpretation of the data. Experimental studies of the underlying event have been performed at the Tevatron [9] but their extrapolation to the LHC energy is rather uncertain [10]. Again, the initial LHC data will be important to perform experimental studies and to improve the theoretical description of this effect.

A new experimental problem at the LHC is the event pile-up due to the high instantenous luminosity. Also these interactions must be, at least on a statistical basis, separated from the hard scattering process which requires experimental studies as well as a very good theoretical description.

1.2.2 QCD and jets

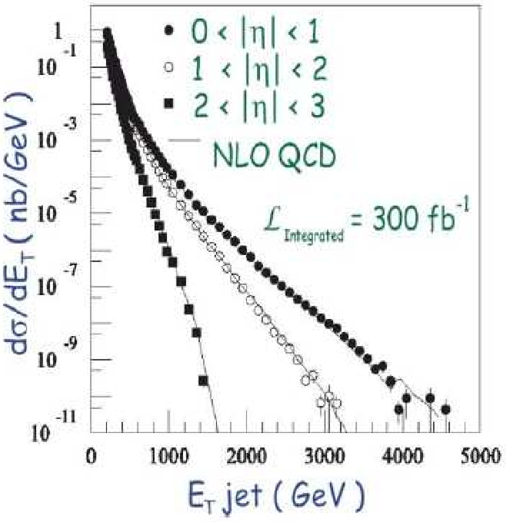


Figure 1.5: *Differential jet cross section as a function of the transverse jet energy (ATLAS).*

Hadron colliders are a predestined tool for probing strong interactions and the physics of jets. The observable spectrum of transverse jet energies (E_T) ranges from about 200 GeV to several TeV with the corresponding cross sections varying over 11 orders of magnitude. This is illustrated in Fig. 1.5, showing the differential cross section $\mathrm{d}\sigma/\mathrm{d}E_T$ in various intervals of rapidity. Of particular interest are the jets at the highest energies. For an integrated luminosity of 300 fb^{-1} which corresponds to a few years of LHC operation at the highest instantaneous luminosity about 4 million jets with $E_T > 1$ TeV and still 400 with $E_T > 3$ TeV are expected. These samples will allow to probe QCD at the smallest distance scales and provide information on a possible quark substructure.

1.2.3 Electroweak physics

The LHC will abundantly produce W and Z bosons which provide relatively clear signatures and are readily identifiable through their leptonic decays. The experiments will hence be able to study the properties of W and Z bosons and to complement and extend the tests of the electroweak theory performed so far at e^+e^- and $p\bar{p}$ colliders.

1.2.3.1 The mass of the $\mathbf{W}^{\pm}$ bosons

An important parameter for precision tests of electroweak theory, which has a fundamental role in the global fit of the Standard Model quantities to estimate the Higgs mass, is the mass of the W boson. Techniques to measure this parameter at hadron colliders have been developed and already today the Tevatron experiments provide results competitive to the precision reached at LEP. At the LHC the overwhelming statistics on Z bosons can be used to experimentally constrain theoretical uncertainties on the production and to determine accurately the detector response. This will minimise the dependence on Monte Carlo simulations and the corresponding uncertainties. A method for tuning the resolution in Monte Carlo simulation will be accurately illustrated in the final chapter of my thesis, as main topic of my analysis.

Fig. 1.6 shows the expected experimental distribution of the electron transverse energy in selected $W \rightarrow e\nu$ decays. The background contributions from other processes are small and already with 1 fb^{-1} of integrated luminosity large statistics can be accumulated. The determination of the W mass from this and other distributions will be limited by experimental and theoretical systematic uncertainties. Detailed studies [11] show that the LHC experiment have the potential to further improve the error on m_W beyond the anticipated uncertainty of about 20 MeV from final LEP and Tevatron Run II data. A very good understanding of the detector and improved theoretical calculation are required to achieve this goal.

1.2.3.2 The Drell-Yan process

Drell-Yan production of lepton pairs $q\bar{q} \rightarrow e^+e^-$, $\mu^+\mu^-$ is the equivalent of the reaction $e^+e^- \rightarrow$ hadrons which was extensively studied at LEP and SLC. This process is not only an important tool to constrain the PDFs but also the discovery channels for heavy new particles decaying into leptons, e.g. a Z' boson.

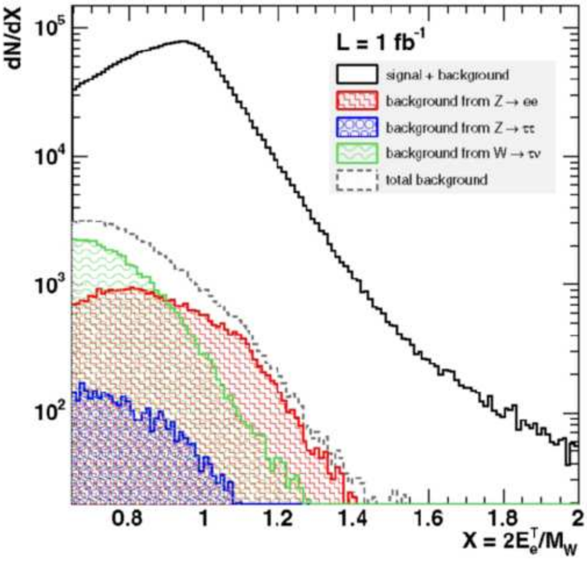


Figure 1.6: *Expected transverse energy spectrum of electrons scaled to the W mass in the channel* $W \rightarrow e\nu$ *together with the contributions from background sources (CMS).*

Fig. 1.7 shows the expected number of events per di-lepton channel and experiment as a function of the mass for a total luminosity of 100 fb^{-1} in comparison to the expectation at the Tevatron Run II (10 fb^{-1}) [12]. The higher luminosity and energy at the LHC will extend the sensitivity well beyond the reach of the Tevatron into the TeV range. Already with the modest integrated luminosity of approximately 1 fb^{-1} the sensitivity of the LHC will extend beyond the reach of the Tevatron experiments.

1.2.4 Physics of the top quark

The LHC is also a real top quark factory which will permit precision measurements of the properties of the heaviest SM fermion. Contrary to the Tevatron the dominant production process for top quark pairs is gluon fusion $gg \rightarrow t\bar{t}$ contributing with about 87% much more than quark annihilation $q\bar{q} \rightarrow t\bar{t}$. The top quark decays to almost 100% into a b quark and a W boson such that the signature of a $t\bar{t}$ pair is given by the decay topology of the W pair in addition to the two b quarks in the final state. The experimentally cleanest modes are characterised by the leptonic (electron or muon) decay of one or both W bosons which are referred to as semi-leptonic and di-lepton channel, respectively.

The experiments will be able to establish a $t\bar{t}$ signal very quickly with relatively low integrated luminosity. As an example Fig. 1.8 shows the result of a selection for di-leptonic events for 1 fb^{-1} [13]. After removing the contribution of Z + jets events with a cut on the invariant mass of

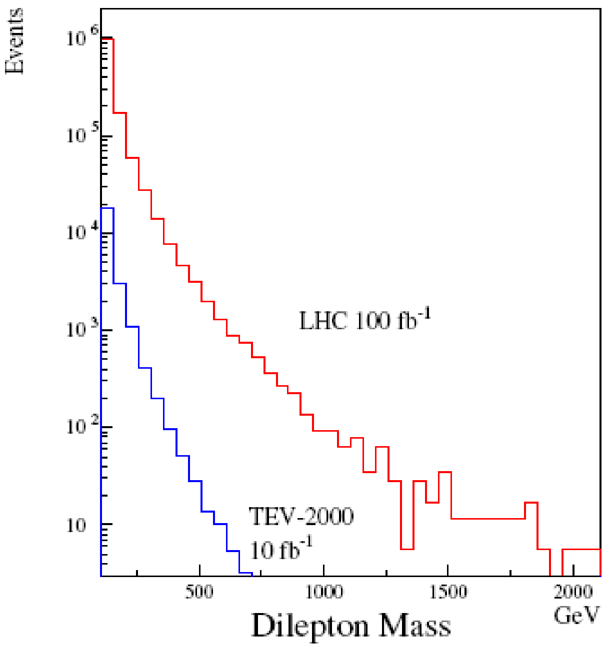


Figure 1.7: *Mass spectrum of Drell-Yan lepton pairs in 50 GeV bins.*

the lepton pair the only significant background remaining are other $t\bar{t}$ final states.

1.2.4.1 Measurement of the top quark mass

An important measurement is the precise determination of the top quark mass which is a crucial quantity for precision tests of the SM as it enters quadratically in quantum corrections. The most promising channel for a precise top mass measurement is the semi-leptonic final state which provides a clear signature from the high momentum lepton together with a high branching ratio and full mass information from the hadronically decaying top quark. The top mass is then reconstructed from the two light quark jets exploiting a W mass constraint and one b-quark jet.

Fig. 1.9 shows a mass distribution obtained from kinematic reconstruction of the semileptonic events [14]. Very clean top samples can be obtained in this channel where the signal to background ratio can be adjusted by the selection criteria. This offers the possibility to optimize the analysis for the available integrated luminosity, balancing statistical and systematic errors and to obtain a significant result for the top mass rather quickly. The achievable ultimate precision from this channel is expected to be $\Delta m_t \approx 1$ GeV$/c^2$ where the knowledge of the energy scale of b quark jets and of final state radiation will be the main limiting factors. Also the other $t\bar{t}$ decay modes will contribute to the top mass determina-

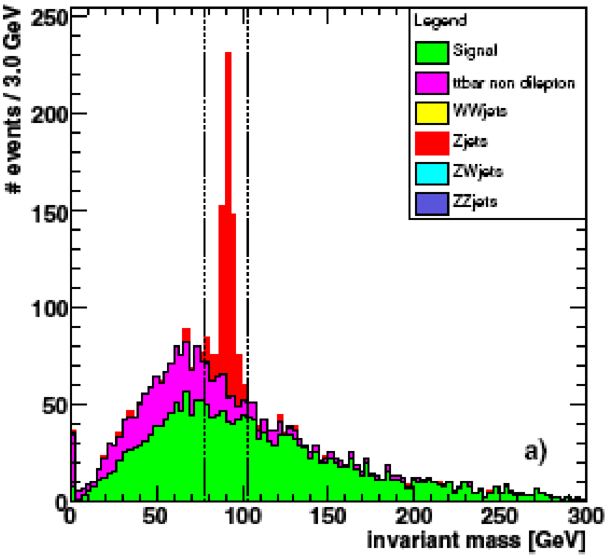


Figure 1.8: *Mass distribution of the lepton pair in di-lepton $t\bar{t}$ candidates corresponding to 1 fb^{-1} (CMS). The broken lines indicate the mass window used to remove Z + jets events.*

tion. The fully hadronic decay mode has a large branching ratio of 45% but the large QCD background and the six jets in the final state create a difficult environment and ambiguities. Even though the two neutrinos present in the di-lepton channel lead to kinematically under constraint events, methods have been developed and applied to exploit this channel for the top mass measurement [13]. Another interesting mode for the mass measurement are top decays where a J/Ψ meson emerges from the b fragmentation and the W decays into an electron or muon. The J/Ψ is readily identifiable by its decay into two leptons and the invariant mass of the J/Ψ and the lepton from the W decay is rather sensitive to m_t [14]. The partially reconstructed top mass is independent of the b jet energy scale, thus circumventing the problem limiting the precision of the inclusive semi-leptonic channel.

1.2.4.2 Total and differential cross sections

Evidently the determination of the total cross section $pp \rightarrow t\bar{t}$ at 14 TeV is an important and interesting measurement in itself. Moreover as it depends strongly on m_t it can provide another handle on the top mass provided that the experimental, namely the luminosity, and theoretical uncertainties can be controlled to the percent level. Also precise measurements of differential cross sections as $\mathrm{d}\sigma_{t\bar{t}}/\mathrm{d}p_T$ and $d\sigma_{t\bar{t}}/\mathrm{d}\eta$ are very valuable as they constrain quark and gluon PDFs at high Q^2 values and

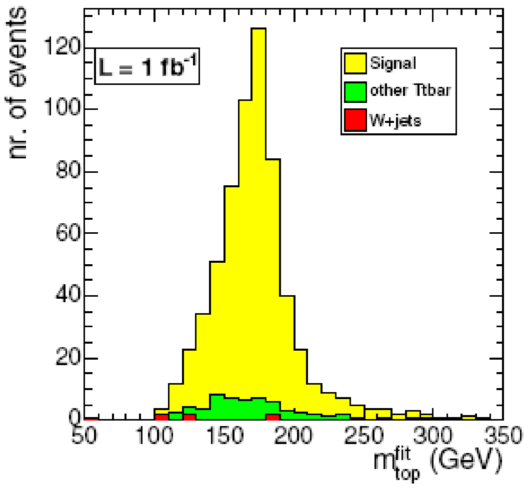


Figure 1.9: *Top mass distribution in the semileptonic channel obtained from a kinematic fit and using 1 fb*$^{-1}$ *(CMS).*

help to reduce uncertainties on discovery channels like Higgs and SUSY production channels.

Of particular interest is also the invariant mass distribution of the $t\bar{t}$ system. A hypothetical particle decaying preferentially into the heaviest SM particle would show up as a resonance in the differential cross section $d\sigma_{t\bar{t}}/\mathrm{d}m_{t\bar{t}}$.

The large statistics available at the LHC will allow to search for rare decays of the top quark. Such analyses comprise rare SM decays such as the CKM suppressed modes $t \to W^+s$, W^+d and the higher order processes $t \to W^+Zb$, $t \to W^+W^-c$ as well as flavour changing neutral current decays like $t \to cZ$.

1.2.5 The Standard Model physics as a mean to calibrate detectors

Beyond constituting the possibility to test the Standard Model to an unprecedent scale, the analysis of some of the processes described until now will allow to calibrate the detectors in the first phase. Exploiting the di-lepton mass resonances (Z, J/ψ and Υ) will permit to keep under control the lepton energy scale, whilst the measurement of the W transverse mass will allow the calibration of the missing transverse energy. Finally, with the huge rate of $t\bar{t}$ it is possible to calibrate jets and measure b-tagging efficiency. Some processes will be very useful at the first stages to tune the Monte Carlo SM description. Moreover W/Z + jets are the largest backgrounds for many new physics searches therefore the inclusive Z p_T spectrum and the rate of events with different jet multiplicities will be measured from data.

Once the main experimental systematics will be under control and the Monte Carlo description of the SM at the LHC energy scale will be tuned from data, the conditions will have been established for any possible claim of new physics.

1.3 The Higgs boson search at the LHC

One of the main goals of the two general purpose experiments at LHC, CMS and ATLAS is to study the mechanism that breaks the symmetry of the SM Lagrangian giving rise to the particle masses. Within the Standard Model, this means to search for the Higgs boson from $m_H =$ 100 GeV to $m_H = 1$ TeV. If the Higgs is found, it is necessary then to understand if it is a SM Higgs or a SUSY Higgs.

If the Higgs is not found, alternative models will have to be considered. In this case, search for new physics will become the main aim.

Concerning supersymmetry, all the s-particles with mass $m_{\tilde{s}} \leq$3 TeV will be accessible. For exotic models (like extra dimensions, composite particles, technicolor, new strong interaction, new lepton families, additional bosons) the mass reach is 5 TeV. The detailed description of the physics beyond Standard Model goes beyond the purpose of this thesis.

1.3.1 The Standard Model Higgs boson

The Higgs boson mass is a free parameter of the Standard Model, which means that it is not predicted by the theory. On the other hand, the Higgs boson couplings to the fermions and bosons are predicted to be proportional to the corresponding particle squared masses. This implies that the Higgs boson production and decay processes are dominated by channels involving the coupling to heavy particles. In particular for m$_H$ $<$ 130 GeV/c^2 the Higgs boson mainly decays in $b\bar{b}$, while for m$_H$ $>$ 130 GeV/c^2 it preferably decays into the vector bosons. Regarding the remaining gauge bosons (photons and gluons), the coupling cannot be at tree level, but only by one-loop graphs where the main contribution is given by *top* loops for the $gg \rightarrow H$ channel and by W^+W^- and *top* loops for the $\gamma\gamma \rightarrow H$ channel.

1.3.1.1 The Higgs Boson production

The main processes contributing to the Higgs boson production at a hadron collider are represented by the Feynman diagrams in Fig. 1.10. The corresponding cross sections for a center of mass energy $\sqrt{s} = 14$ TeV are shown in Fig. 1.11.

- *gluon-gluon fusion*

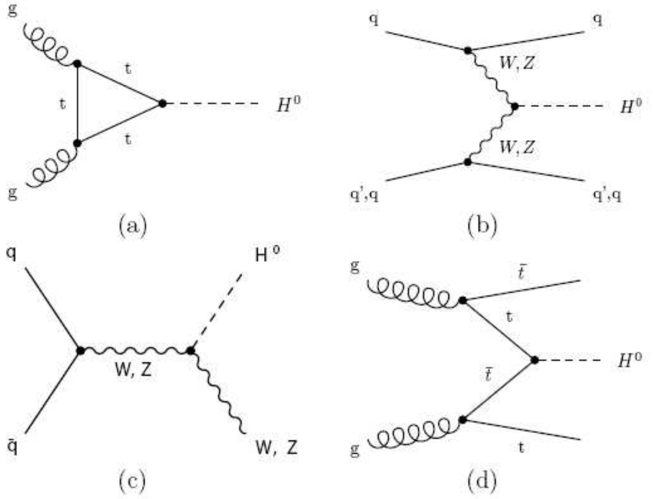


Figure 1.10: *Typical diagrams for all relevant Higgs boson production mechanisms at leading order: (a) gluon fusion, (b) vector boson fusion, (c)* Higgs-strahlung*, (d) Higgs bremsstrahlung off top quarks.*

The *gg* fusion, shown in Fig. 1.10, is mediated by top and bottom quark loops (Fig. 1.10a). Due to the large size of the Yukawa couplings and the gluon densities, this process will provide the largest production cross section at the LHC over the whole Higgs boson mass spectrum. A 50 to 100% increase of the cross section is to take into account if the next-to-leading (NLO) QCD corrections to the top and bottom quark loops [17, 18, 19] are considered in the calculation. Another $\sim$ 20% increase in the total cross section shows up in the limit of very heavy top quarks, for which the NNLO QCD corrections have been recently calculated [20, 21, 22]. A full massive NNLO calculation is not available but the approximate NNLO results have been improved by a soft-gluon resummation at the next-to-next-to-leading log (NNLL) level, which yields another increase of the total cross section by about 10% [23]. Electroweak corrections have been computed and turn out to be small [24, 25, 26, 27]. The theoretical uncertainties of the total cross section can be estimated as 20% at NNLO due to the residual scale dependence, the uncertainties of the parton densities and due to neglected quark mass effects. The production of the Higgs boson through gluon fusion is sensitive to a fourth generation of quarks. Finally, as the Higgs boson couples proportionally to the fermion mass, including a fourth generation of very heavy quarks will more than double the cross section.

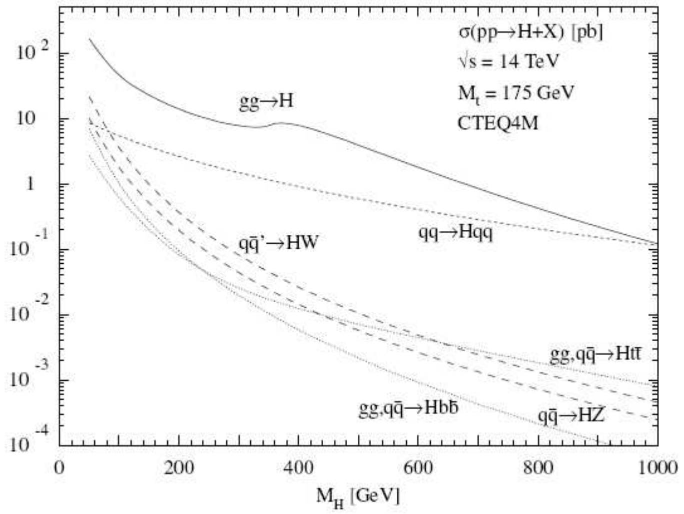


Figure 1.11: *Higgs boson production cross sections at* $\sqrt{s} = 14$ *TeV as a function of the Higgs boson mass. The cross sections are calculated using HIGLU and other programs [15]; they contain higher order corrections and the CTEQ6m [16] PDF has been adopted.*

- *vector boson fusion*

  For large Higgs masses the W and Z boson-fusion processes (Fig. 1.10b) become competitive. These processes are relevant in the intermediate Higgs mass range, too, since additional forward jets offer the opportunity to reduce background processes significantly, providing a clear experimental signature. Both leading and next-to-leading order cross sections for this process are known with small uncertainties and the higher order QCD corrections are quite small.

- *higgs-strahlung*

  In the intermediate mass range $M_H < 2M_Z$ Higgs-strahlung off W, Z gauge bosons (Fig. 1.10c) provides alternative signatures for the Higgs boson search. The cross section for this process is several order of magnitudes lower than gg fusion and VV fusion ones. Since only the initial state quarks are strongly interacting at LO, the NLO QCD corrections can be inferred from the Drell-Yan process. They increase the total cross section by $\mathcal{O}(30\%)$ [28, 29, 30]. The NNLO has been recently calculated and are resulted to be small [31]. Moreover, the full electroweak corrections have been calculated, resulting in [32] a decrease of the total cross sections by 5 - 10%. The total theoretical uncertainty is of O(5%).

- *radiation off top quark*

  This channel (Fig. 1.10d) plays a significant role for Higgs masses below ~ 150 GeV/c^2. Also in this case the cross section is several order of magnitudes lower than the ones of gg and VV fusion, but a good experimental signature is provided by the presence of the $t\bar{t}$ pair in the final state. The LO cross section in known since long time ago [33, 34], while the full NLO QCD calculations are more recent and resulted in a moderate increase of the cross section by ~ 20% [35, 36].

1.3.1.2 The Higgs Boson decay

The Higgs decay modes can be divided into two different mass ranges and are shown in Fig. 1.12. For M_H lower than ~ 135 GeV/c^2 the Higgs boson mainly decays into $b\bar{b}$ and $\tau^+\tau^-$ pairs with branching ratios respectively of about 85% and 8% (Fig. 1.12, right plot). The decay modes into $c\bar{c}$ and gluon pairs, with the latter mediated by a top and bottom quark loops, accumulate branching ratio of up to about 10%, but do not play a relevant role at the LHC. One of the most important Higgs decays in this mass range at the LHC is the decay into photon pairs, mediated by W boson, top and bottom quark loops. It reaches a branching fraction of up to 2×10^{-3}. The NLO QCD [37, 38] and electroweak [39, 40, 41] corrections are known to be small in the Higgs mass range relevant for LHC.

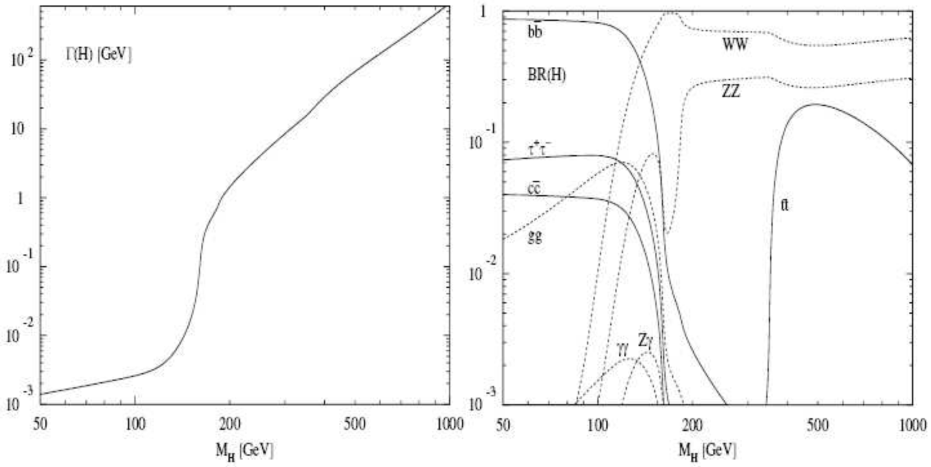


Figure 1.12: *Left plot: total decay width (in GeV/c^2) of the SM Higgs boson as function of its mass. Right plot: branching ratios of the dominant decay modes of the SM Higgs particle. All relevant higher-order corrections are taken into account.*

For Higgs masses above 135 GeV/c^2 the main decay modes are those into WW and ZZ pairs, where one of the vector bosons is off-shell below the corresponding kinematical threshold. These decay modes dominate over the decay in $t\bar{t}$ pairs, whose branching ratio does not exceed $\sim 20\%$. The electroweak corrections to the vector bosons decay modes are of moderate size [42, 43, 44, 45].

The total decay width of the Higgs boson, shown in Fig. 1.12 (left plot), does not exceed about 1 GeV/c^2 below the WW threshold. For very large Higgs masses the total decay width grows up to the order of the Higgs mass itself so that the interpretation of the Higgs boson as a resonance becomes questionable.

1.4 The CMS Experiment

CMS is one of the two general purpose experiments at LHC. Its structure is driven by the configuration of the magnetic field, which has been chosen in order to maximize the BL^2 term to optimize the resolution on the measurement of the muon momentum. In fact, in order to fulfill its physics program, a good resolution for muons from few GeV up to 1 TeV is mandatory.

The interesting particles emerging from the $p-p$ collisions are produced over a wide range of energy (from few hundred MeV to a few TeV) and over the full solid angle. The ideal detector would not permit to any particle to escape unseen. This would imply to avoid any crack in the acceptance, which is not possible of course. It is nevertheless possible to minimize the cracks.

The CMS experiment has chosen a *compact* layout with a solenoid producing a very intense magnetic field $B = 3.8$ T in a moderate volume ($R = 3$). The main technological challenge for CMS is to reach such a high magnetic field uniformly over a large volume. This is obtained with a superconducting solenoid which encloses the three innermost detectors (the inner tracker and the two calorimeters).

Briefly, the structure of the CMS detector (Fig. 1.13), going from the innermost detector towards outside, is the following (more details can be found in [51]):

- **Inner Tracker**

  The CMS tracker is completely made of silicon strips and pixels sensors. The resolution on the charged particle momentum is $\sigma_{pT}/p_T \sim 1.5 \times 10^{-4} p_T \oplus 0.005$. This detector will be described in more detail in Chapter 2.

- **Electromagnetic calorimeter**

  The homogenous electromagnetic calorimeter is made of $PbWO_4$ crystals. Measuring the energy resolution with the electron beams

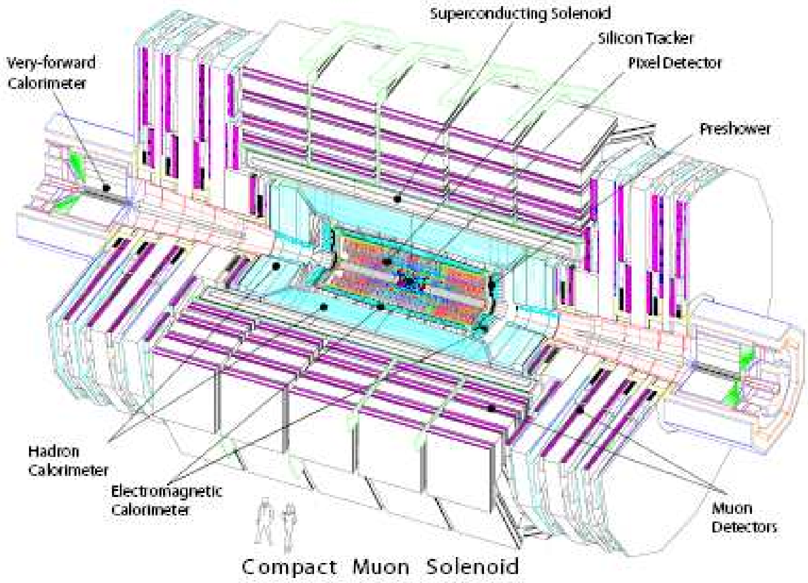


Figure 1.13: General view of the CMS detector.

with momenta between 20 and 250 GeV and limited to a 4 × 4 mm^2 region around the point of maximum containment of the tested supermodule, the value obtained is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right) + \left(\frac{0.12}{E}\right)^2 + (0.30\%), \tag{1.6}$$

with E in GeV.

- **Hadronic calorimeter**

  The CMS hadronic calorimeter is made of Cu-scintillator with an energy resolution of

  $$\frac{\sigma_E}{E} \sim \frac{100\%}{\sqrt{E}} \oplus 0.05, \tag{1.7}$$

  where E is in GeV. As the hadronic calorimeter is located inside the magnet and so it is limited from its size, it is not long enough to contain the hadronic showers entirely. For this reason an additional tail catcher (the HO detector, a layer of scintillators) has been placed after the calorimeter in order to limit the punch through into the muon system.

- **Muon detectors**

  Muons provide clean and unambiguous signatures of much of the physics that CMS was designed to study. The muon system is composed of Resistive Plate Chambers (RPC) and Drift Tube chambers (DT) in the barrel, and of DT and Cathod Strip Chambers (CSC) in the endcaps. All the chambers are placed in the iron of the magnet yoke and the muon transverse momentum resolution is $\sigma_{pT}/p_T \sim 5\%$ at 1 TeV.

- **CMS Trigger**

  For the nominal LHC design luminosity of 10^{34} cm^{-2}s^{-1}, an average of 17 events occurs at the beam crossing frequency of 25 ns, a time interval too small even to read out all raw data from the detector. This input rate of 10^9 interactions every second must be reduced by a factor of at least 10^7 in order to achieve a frequency of 100 Hz, the maximum rate that can be archived by the on-line computer farm. CMS has chosen to reduce this rate in two steps, namely Level-1 Trigger (L1) [52] and High Level Trigger (HLT) [53]. At the first level all data are stored for 3.2 μs, after which no more than 100 kHz of the stored events are forwarded to the High Level Triggers. This must be done for all channels without dead time. The HLT then can operate on longer timescales, but always consistent with the overall Level-1 output rate. The L1 system is based on custom electronics, while the HLT system, relies upon commercial processors. The former uses only coarsely segmented data from calorimeter and muon detectors, while holding all the high-resolution data in pipeline memories in the front-end electronics. The latter is provided by a subset of the on-line processor farm which, in turn, passes a fraction of these events to the remainder of the on-line farm for more complete processing.

  The trigger is the start of the physics event selection process. A decision to retain an event for further consideration has to be made every 25 ns. This decision is based on the event's suitability for inclusion in one of the various data sets to be used for analysis. The data sets to be taken are determined by CMS physics priorities as a whole. These data sets include di-lepton and multi-lepton data sets for top and Higgs searches, lepton plus jet data sets for top physics, and inclusive electron data sets for calorimeter calibrations. In addition, other samples are necessary for measuring efficiencies in event selection and studying backgrounds. The trigger has to select these samples in real time along with the main data samples.

Chapter 2

The CMS Tracker: Structure, Calibration and Local Reconstruction

The detection and study of the diverse signatures for new physics at the LHC will rely on the clean identification and precise measurement of leptons, photons and jets. Robust tracking and detailed vertex reconstruction within a strong magnetic field are essential tools to reach these objectives. In order to meet these requirements, the CMS Collaboration proposed to adopt a Silicon Pixel vertex detector close to the beam pipe surrounded by a Silicon Strip Tracker (SST). Its goal is to provide a precise momentum estimate for charged particles, and to allow a precise determination of the position of secondary vertices.

LHC events will be very complex, hence track reconstruction comes as a complex pattern recognition problem. In order to ease pattern recognition two requirements are fundamental:

- low detector occupancy,
- large hit redundancy.

The low hit occupancy is achieved with a highly granular detector, while the redundancy is realized with a large number of detecting layers.

In this chapter the CMS Tracker [54] is described in detail.

In Sec. 2.1 a brief overview of the Pixel Vertex Detector is given. Then starting from Sec. 2.2 onward the Silicon Strip Tracker is described in all the details, from the operation principle of the single detection unit to the layout and readout system, from the calibration to the principles of the reconstruction algorithm and of the alignment procedures.

2.1 The Pixel Vertex Detector

The CMS pixel detector shown in Fig. 2.1 is the innermost part of the tracking system. It has a paramount importance as a starting point in the reconstruction of the charged particle tracks. It covers a region of pseudorapidity $|\eta| < 2.4$ and it is constituted of three barrel layers, each 53 cm long, positioned at $r = 4.4$, 7.3 and 10.2 cm, and two disks per each side, placed at $z = \pm 34.5$ cm and $z = \pm 46.5$ covering radii between 6 and 15 cm to guarantee at least two crossed layers per track. When the high luminosity condition will be reached, the inner barrel layer will be substituted by an outer layer placed at $r = 13$ cm to improve the resolution and limit the radiation damages.

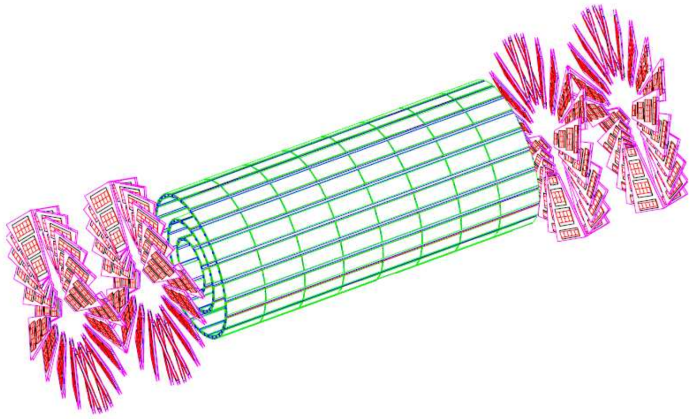


Figure 2.1: *Perspective view of the CMS pixel system in the high luminosity configuration.*

Each layer is composed of modular detector units, containing a 250 μm thin-segmented sensor plate with highly integrated readout chips connected to each pixel using bump-bonding technique [55]. A scheme of a pixel detector unit is shown in Fig. 2.2. A squared pixel shape has been chosen in order to optimize measurements in $r\phi$ and z coordinates, both important for vertex finding and impact parameter resolution. The pixels have a size of 150×150 μm^2 and are combined with analog signal readout to profit of charge sharing effects among pixels and improve position resolution by interpolation.

The charge sharing between pixels is due to the Lorentz drift of charge carriers, which is about 32° for electrons in a 4 T magnetic field, three times wider than for the holes. Therefore initial n-type substrate sensors are chosen to collect electron signals on n^+ implants, which in turn are more radiation hard. However it can not be avoided that in the barrel layers with $r < 10$ cm the pixel and the readout chip lifetime is reduced by hostile radiation environment below CMS lifetime. Hence the layer at $r = 7.3$ cm should be replaced after six or seven years of operations.

In the barrel the pixels are tilted to induce significant charge sharing between neighboring implants in $r\phi$ plane improving the intrinsic *hit* resolution down to 10-15 μm, far below the 150 μm width of each n^+ implant. Charge sharing is present also along z direction for inclined

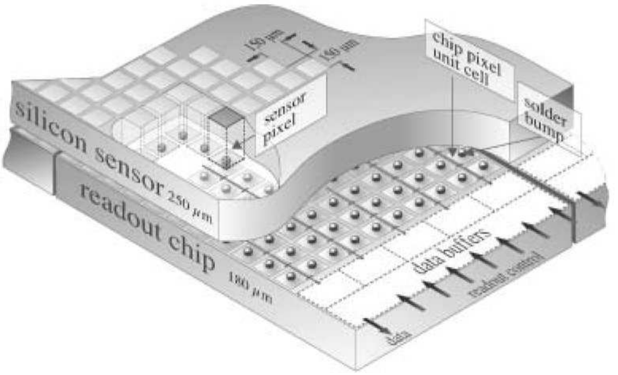


Figure 2.2: *Schematic view of a pixel detector unit.*

tracks, leading to a similar resolution.

The detectors placed on the disks are rotated with an angle of 20° around the central radial axis to benefit of charge sharing improved both in r and $r\phi$ directions by induced Lorentz effects. Despite a reduced Lorentz angle with respect to the barrel case, the resolution in r and $r\phi$ is expected to be 15 μm at CMS start and degrading to 20 μm when radiation damages arise.

Each pixel signal is readout by a Pixel Unit Cell (PUC) bump-bounded directly to the pixel module. The PUC is integrated on the readout chip, which attends to 52 columns and 80 rows for a total of 4160 pixels. Since the number of channels is very high (44 millions), zero-suppression is mandatory to reduce the huge data volume down to a reasonable size. To perform the Zero Suppression each PUC is equipped with an analog circuit, which provides a logical positive output if the collected signal exceeds a tunable threshold. To reduce the number of channels to readout, two near-by PUC columns are read by one circuit placed in the periphery. The analog signals are temporarily stored into dedicated pipelines and on positive Level-1 trigger decision are transmitted through optical fibres to the front-end driver in the counting room.

2.2 The Silicon Strip Tracker

2.2.1 Silicon strip detectors working principle

The outer part of the tracker is made with layers of silicon microstrip detectors. Its fundamental active element is the *module*, which is made of a silicon microstrip sensor together with its associated readout electronics. The sensor is an n-type phosphorus doped substrate with p^+ implant strips, as shown in Fig. 2.3.

The p^+-n junction is reversely biased applying a positive voltage (hundreds of Volts) to the n side keeping the strips at ground. In this way the region between the junction and the backplane is completely depleted of free charge carriers, with exception of the termally created ones. When an ionizing particle passes through the silicon, it interacts in the bulk creating electron-hole pairs (e/h) which drift in the electric field towards the backplane and p^+ implants respectively. The mean energy required to create a e/h pair in silicon is 3.6 eV, therefore a minimum ionizing particle ($m.i.p.$) with an average energy loss per path length of 390 eV/μm should create 32500 e/h pairs passing through a 300 μm thick sensor. Since the energy loss distribution is Landau shaped, the most probable value, being about 288 eV/μm, differs from the average. For this reason it is preferred to quote the most probable value of e/h pairs produced by a $m.i.p.$, which is 24000 in a 300 μm sensor and represents a rough estimate of the charge to collect.

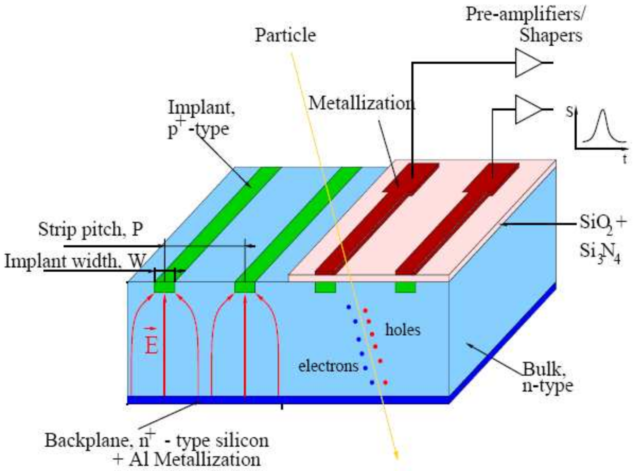


Figure 2.3: *Principle of operation of a silicon microstrip detector.*

Some arrangements are made to permit a better working of the devise. The n^+ implant backplane forms an ohmic contact and avoids charge injection into the bulk from the connected aluminium layer. On the opposite side the active area is surrounded by two p^+ implants. The inner is a *bias ring* and is used to uniformly bias the strips through 1.5 MΩ polysilicon resistors, the outer is a *guard ring* and limits the dark current contribution from sensor bounds. At the detector edges, n^+ implants are placed to limit charge injection from the regions damaged by the cut on the wafer. Insulating capacitor layers of dielectrics (SiO_2, Si_3N_4) are grown between the p^+ implants and the aluminum strip electrodes to decouple the readout electronics from the detector leakage current.

The strips are bonded to an array of readout chips APV [56] housed on a thin hybrid circuit. The analog signal of each strip is transmitted to ADCs located in the counting room via optical links.

With these microstrip sensors it is possible to measure one coordinate, interpolating the crossing position by means of the charge sharing between adjacent strips. A configuration to allow the bidimensional measurement of coordinates is realized gluing two detectors back-to-back with 100 mrad tilted strip directions. This configuration is referred to as *stereo* or double-sided and is preferred to the pixel segmentation, although the resolution is worse, because the number of readout channels is smaller. The *hit* ambiguities present in this detector configuration are resolved at track reconstruction level.

2.2.2 The Tracker layout

The silicon strip tracker covers a tracking volume up to $r = 1.1$ m with a length of 5.4 m and is divided in three parts:

- **inner tracker**

  4 barrel layers (Tracker Inner Barrel - TIB)

  3 disks per endcap (Tracker Inner Disks - TID)

- **outer tracker**

  6 barrel layers (Tracker Outer Barrel - TOB)

- **tracker endcaps**

  9 disks per endcap (Tracker End-Cap - TEC)

The TIB has four layers assembled in shells; the two innermost layers host double-sided detectors, pointed out in blue in Fig. 2.4. The two TIDs, each one made of three small disks, complement the TIB region in the forward part. The outer barrel structure (TOB) consists of six concentric layers, and also in this case the two innermost are double-sided. The TEC modules are mounted on nine disks on both sides of the barrel. The disks are then subdivided in rings, with trapezoidal-shaped modules to follow the ring geometry [57]. The detectors of ring 1, 2 and 5 are made of double-sided modules.

The main differences between the inner and the outer tracker are the thickness and dimensions of the silicon modules. The inner tracker is made with *thin* sensors 320 μm thick, with 117 mm long strips and 64 mm total width. The outermost modules have *thick* sensors, the bulk thickness being 500 μm, with 190 mm long strips (because of the lower occupancy) and total width of 96 mm. The layer thickness permits to collect a larger signal to compensate the higher noise due to longer strips. Moreover, this size reduces both the front-end electronic channels and the sensor costs, because 500 μm thick sensors are produced by 6" wafer commercial lines in a cheapest way.

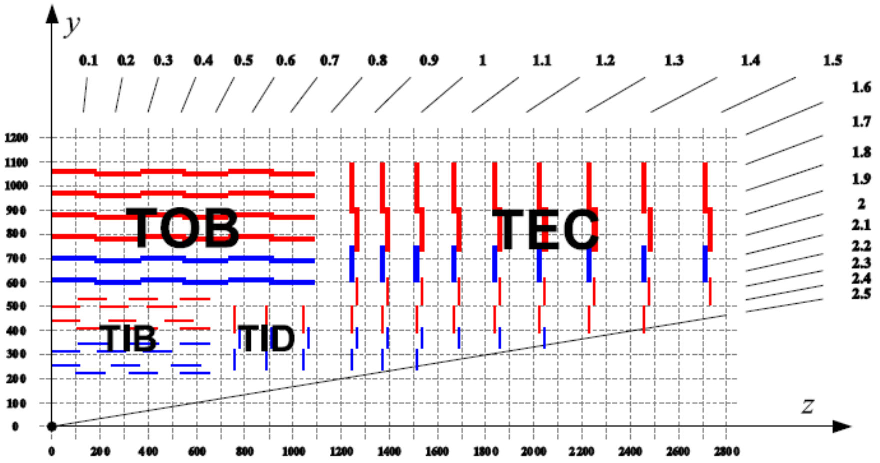


Figure 2.4: *View of a quarter of the silicon strip tracker in the rz plane. The horizontal and vertical axes are in mm; the top axis shows the pseudorapidity. The red lines represent single-sided module layers, blue lines stereo (double-sided, back-to-back) module layers. Modules marked with a thin (thick) line are equipped with 320 μm (500 μm) thick sensors.*

Tracker	detectors	thickness [μm]	pitch [μm]
Inner Barrel (TIB)	2724	320	81/118
Outer Barrel (TOB)	5208	500	123/183
Inner Disks (TID)	816	300	97/128/143
Outer Disks (*thin*) (TEC)	2512	300	96/126/128/143
Outer Disks (*thick*) (TEC)	3888	500	143/158/183

Table 2.1: *Thicknesses and pitches of the tracker sensors in the different substructures.*

The TEC modules are divided in two categories: thick substrates for the outermost three rings, thin for the rest. In Tab. 2.1 the thickness and the distance between two adjacent strips (the *pitch*) for the tracker sensors are reported. In the barrel, the modules are rectangular-shaped, with the strips parallel to the beam direction for φ and r coordinate measurements. The endcap modules are trapezoidal-shaped (wedge-shaped) to allow a radial strip disposition for φ and z measurements. The total number of modules is 15 148 (about 198 m^2 silicon active area), 6054 *thin* and 9 096 *thick* for a total of 9 648 128 electronic channels (strips) to be readout in groups of 128 by 73 736 APV chips [58].

2.2.3 The readout system

The strip signal is collected through the metallized strip by one of the 128 input channels of the APV chips placed on the front-end hybrid. Since the strip pitch is different from module to module and only one type of readout chip is foreseen, a *pitch adapter* is designed to connect groups of 128 strips to the input pads of the APV. The signal of each strip is readout through a charge sensitive amplifier followed by a RC-CR shaper with a time constant of about 50 ns. The output voltage is sampled at 40 MHz, synchronously with the LHC bunch crossing frequency, and analog signals are stored temporarily into a pipeline. The pipeline is a 128×192 matrix of capacitor cells that contain the output of all the 128 strips for a maximum of 192 locations, which correspond to the number of bunch crossings and amount to a total time depth of 4.8 μs. The dimension of a cell is 30×35 μm^2 and comprises two transistors to perform the read and write operations, and a 0.25 pF storage capacitor.

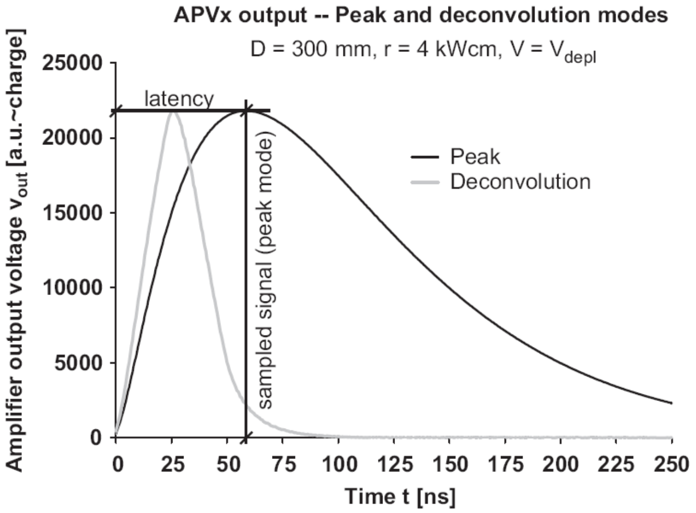


Figure 2.5: *Output signal of the APV25 amplifier in peak and deconvolution mode. Only in deconvolution mode a proper separation of individual bunch crossings can be achieved for an LHC like timing of 25 ns between two events. The curve is sampled at a given time (latency) and the sampled signal height is stored. Careful adjustment of the latency setting is important to sample the signal curve at the maximum.*

In case of Level-1 positive response, the pulse height signals stored into the pipeline are processed by an Analog Pulse Signal Processor circuit (APSP), which performs the analysis of the signal samplings in two different scenarios:

- **Peak Mode**: the peak amplitude of the signal is obtained by the shaper output corresponding to the triggered bunch crossing.

- **Deconvolution Mode**: the triggered and the two preceding samples are weighted and combined together to effectively reduce the signal duration to one bunch crossing, at expense of increasing noise[59]. This mode of operation is the most suitable at LHC rates, especially at high luminosity, because it reduces the signal tails of adjacent bunch crossings.

When the trigger is sent to the APV, its output is read. The output "frame" consists of 12 digital control bits followed by 128 analog signals as output of a 20 MHz multiplexer. The total readout time is $(12+128)\times50$ ns $= 7$ μs, hence if two trigger signals were closer than 7 μs, the information from the APV would be lost. To avoid this potentially long APV dead time, an internal FIFO is able to temporary store up to a maximum of 31(10) trigger "frames" when working in peak (deconvolution) mode, thus allowing to absorb the Level-1 trigger time fluctuations.

The pulse height data coming from two APVs are amplified, converted to light by an Analog Opto Hybrid (AOH) and sent via optical fibers to the Front End Driver board (FED), where they are digitized and further processed prior to transmission to the central Data Acquisition (DAQ) system.

The scheme of the readout system for the CMS silicon tracker is sketched in Fig. 2.6.

Clock, trigger signals and slow control communications with the front-end electronics are managed by the Front End Controller (FEC) boards and sent via optical fibers to the Digital Opto Hybrid (DOH) for each control ring of the Tracker. Signals are distributed by the DOH to every Communication Control Unit (CCU) in a control ring. Finally each CCU sends signals to Tracker modules, in particular clocks and triggers via a Phase Locked Loop (PLL) circuit on each module. It receives monitoring data provided by each module from the Detector Control Unit (DCU) chip.

The design of the CMS tracker readout system of Fig. 2.6 is organized to maintain analog data as long as possible basically to interpolate the strip signals and increase the resolution. This should also reduce the complexity of the front-end chips and lower the power dissipation within the tracker volume. The tracker electronics located on the detector is completely optically decoupled from the digitizing and control systems placed in the counting room. To fully exploit this feature in keeping the noise to an acceptable level, also the power supply system (both low and high voltages) is completely floating with a high coupling impedance to the external world in the whole interesting frequency range.

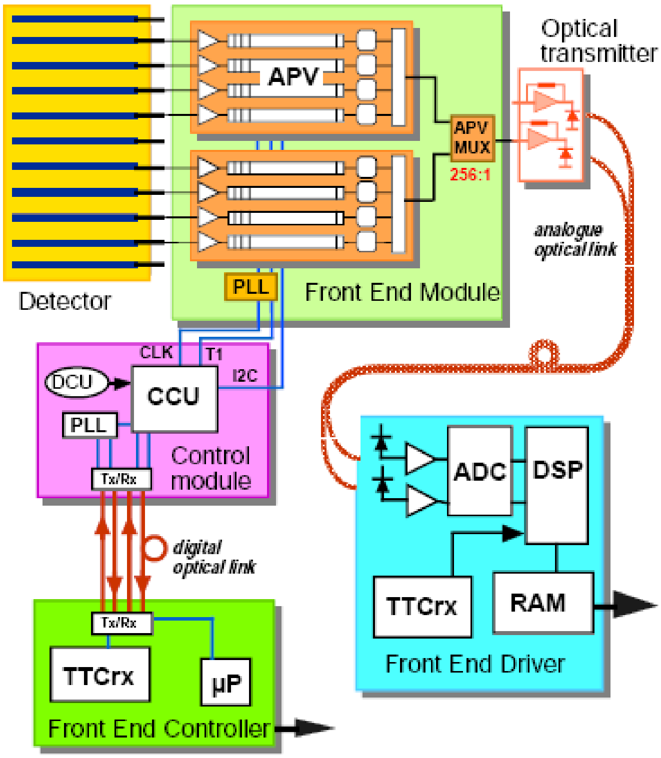


Figure 2.6: *Scheme of the readout system for the CMS silicon strip tracker.*

2.2.4 Radiation damage of silicon microstrip detectors

The tracker is the CMS subdetector closest to the interaction point, hence it has to sustain the higher radiation flux, which deteriorates the modules. The radiation damage is caused not only by the particles produced in primary proton-proton collision, but also by albedo neutrons emitted from the calorimeters surrounding the tracker. Two different effects of the radiation have been observed in silicon detectors: surface damages and bulk damages.

The surface damages [60] are caused by the missed recombination of the e/h pairs created onto the oxide layers of the detector surface by the crossing particles. The macroscopic consequence of this effect is the increasing of interstrip capacitance leading to a worsening of the Signal-to-Noise ratio of the device. A significant increase has been found in the interstrip capacitance after irradiation on detectors with $\langle 111 \rangle$ crystal orientation, but not on sensors with $\langle 100 \rangle$ crystal orientation, therefore the latter have been chosen for the tracker [61].

The interstrip capacitance can be reduced overdepleting the junction, choosing, when possible, a working point around twice the value of the depletion voltage.

Bulk damages originate from the removal of some atoms from their regular sites on the silicon lattice [62]. In this case a point-like defect in the lattice (*vacancy*) is created and can be considered an acceptor impurity, while the displaced atom behaves as a donor; this couple of defects is referred to as *Frenkel pair*. Various combinations and even clusters of these defects can be observed. One of the major effects of bulk damages due to radiation is the change in the effective doping concentration of the silicon crystals, which depends on the fluence of irradiating particles, the time elapsed after exposition and the absolute silicon temperature. The current passing through the junction (*dark current*) is also increased proportionally to the fluence, with the proportionality constant depending on the temperature. As a result, the electronic noise contribution coming from the leakage is enhanced. By lowering the working temperature to -10° C or less this effect is minimized. Furthermore the signal collected to the electrodes is reduced by charge trapping into the damaged bulk. An empirical model, called *Hamburg model* [63], in agreement with experimental data, describes the behavior of the effective doping concentration as a function of the fluence, the annealing time and the storage temperature. When the fluence, normalized to a 1 MeV neutron equivalent, has reached a certain value, the effective bulk donor density approaches zero. At that point, continuing the irradiation, the bulk behaves as p-type

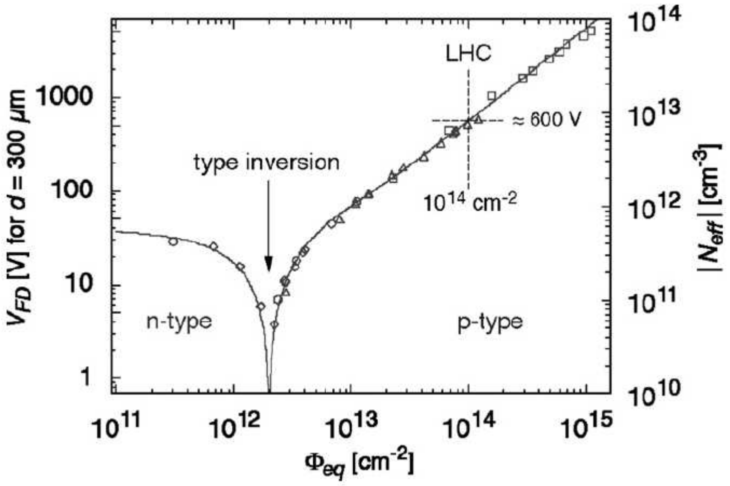


Figure 2.7: *Change in the depletion voltage with respect to the absolute effective doping concentration as measured immediately after the radiation.*

and an effective acceptor density starts to grow up. This phenomenon is called "bulk type inversion" [64] and leads to an increase of the depletion voltages for the highly irradiated silicon sensors, as shown in Fig. 2.7. The reverse bias polarity of the device does not change after the inversion, while the junction moves from the p^+ strip side to the n^+ side.

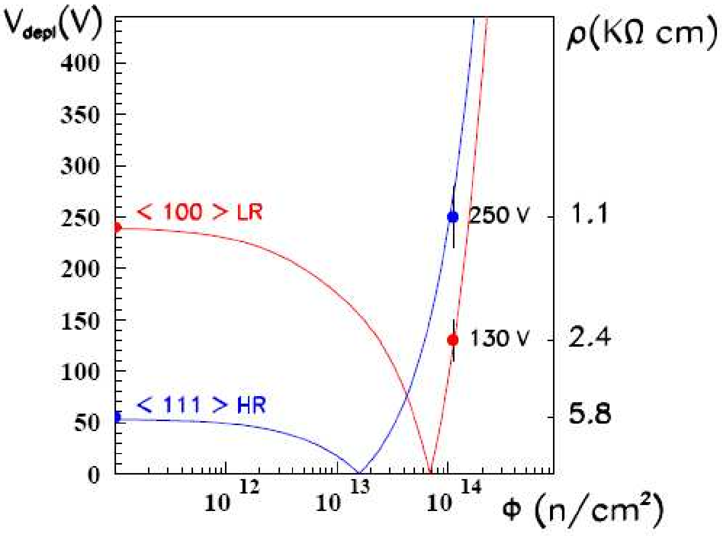


Figure 2.8: *Depletion voltage* V_{depl} *as a function of the neutron fluence* ϕ *for two kinds of silicon detectors,* $\langle 100 \rangle$ *"Low Resisitivity"* $\rho = 1.13 \pm 0.16$ *k*Ω *cm and* $\langle 111 \rangle$ *"High Resistivity"* $\rho = 5.8 \pm 1.1$ *k*Ω *cm.*

In Fig. 2.8 the depletion voltage is shown as a function of the neutron fluence ϕ for two kinds of silicon detectors, $\langle 100 \rangle$ "Low Resisitivity" and $\langle 111 \rangle$ "High Resistivity". The CMS collaboration chose to build the tracker with an initial low resistivity n-type substrate in order to delay the type inversion and keep the detector depletion voltage as low as possible when it is heavily irradiated.

Irradiated detectors should work at low temperature to minimize both the dark current and the *reverse annealing* effect that increase the depletion voltage when the irradiated material is kept at room temperature for a long time. Furthermore, the high bulk current present in irradiated detectors prevents the possibility to operate them through the so called mechanism of *thermal runaway*. This is a positive feedback process that induces an increase of the sensor temperature that can not be removed by the power supply unit. This would lead to an unstable situation making impossible to deplete the detector.

2.3 Commissioning and calibration procedures

Before starting to collect physics data, the detector needs to be fully commissioned and calibrated in order to have the most reliable measurements to be used for physics studies.

The commissioning of the tracker is made in two stages. First the electronic of each module is tuned, adjusting the clock arrival (timing), the AOH gain and bias (Optical Scan), the analog baseline level (VPSP scan), the pulse shape (Calibration scan) and eventually measuring the pedestal and the noise of each strip together with a calibration pulse. In a second stage, a synchronization with the trigger is necessary in order to read the signal from the APV at the right time, at the peak of the signal shape. This is done with a procedure called Latency Scan.

The performance of the tracker strongly depends on the front end electronics supply voltage and its operating temperature. These quantities are measured by the DCU chip located on each module hybrid, which can be read out digitally via the control ring protocol through the electronic chain (DCU, CCU, DOH, FEC). As this electronic chain can be one of the major sources of noise and miscalibration, it is necessary to fully test it and keep every possible effect under control before starting the data taking.

For the SST the calibration procedure can be summarized in the following steps:

- electronic gain measurements
- noise performance studies
- detector quality investigation.

In this section I will just describe the procedures, whilst in the next chapter I will report the results of these calibration during the cosmic data taking in which I have taken part.

2.3.1 Electronic gain measurements

When reading out the signal from the silicon strip modules it is fundamental to have the same gain among the different APVs associated to the same module. One of the calibration procedures is the adjustment of the gains of individual APVs to a common and known value.

Instabilities in the low or high voltages or changes in temperature can also affect the gain. Measuring the electrical stability of the tracker as a function of time, voltage, temperature and other variables should lead to an improved understanding of the likely performance of the full tracker system during actual LHC operations.

The commissioning procedure [76] determines optimal settings of the electronics to achieve a uniform electronic gain. This is obtained by measuring the value of the height of synchronization pulses, referred to as *tick-marks*, generated by the APV. This provides a stable value that can be used as reference to equalize the response of the full electronic chain. The tick-mark height of each APV is required to be within the FED dynamic range and, by varying the AOH offset and gain, should correspond closely to the target value of 640 ADC channels. The full commissioning procedure here described has to be consistently applied whenever changes in coolant temperatures or in the APV parameters or in the hardware configuration are applied. Commissioning runs, referred to as timing runs, provide precise measurements of the tick-mark heights, therefore of the electronic gain.

However, even after the commissioning procedure the tick-mark height distribution can still show a spread in electrical gain due to the coarse precision of the AOH laser gain settings. An offline calibration of the electronic gain is hence necessary to improve the precision of the measurement of noise and signal. This can be achieved by normalizing to an assumed digital header of 640 ADC counts, therefore applying a correction factor of

$$C_{corr} = \frac{640}{TickMark}. \tag{2.1}$$

2.3.2 Noise performance studies

Before irradiation, the noise of a module is almost completely determined by the input capacitance load of the APV chips, which in turn is dominated by the silicon strips. Thus one expects a linear dependence of the noise on the length of the silicon strips for all modules.

Moreover, when the modules are mounted on the final support structures and therefore they are in close proximity to each other, other possible sources of noise can arise. In particular grounding loops, cross talk from neighboring modules, digital noise, cables, power supply, could affect the final noise performance.

The pedestal and noise are measured per strip using a cyclic trigger at low frequency. Pedestals are first calculated as the mean of the raw data and then subtracted from the raw data values. A common mode bias evaluation per chip and per event is then made calculating the median of the pedestal-subtracted data for each APV for each event. This bias is subtracted and then the final pedestal and noise can be calculated as the mean and rms of the final distribution per strip integrated over all events.

The commissioning procedure plays a fundamental role, since it optimizes many system parameters and consequences are observed on the noise performance. For this reason the noise has to be renormalized for each pair of adjacent APV, belonging to the same AOH laser, in order to take into account the different electronic gains. This is done by multiplying the noise for a correction factor.

2.3.3 Detector quality analysis

Complementary to the noise measurement one can study the normality of the noise distribution in order to remove strips with non-Gaussian noise behavior that may lead to unwanted higher occupancy after zero suppression.

For each strip an histogram of the signal after the subtraction of pedestal and common mode noise is filled and fitted to a normal distribution. In order to tag non-Gaussian behavior a Kolmogorov test is performed on the distribution tails with respect to a normal distribution whose central value and width are equal respectively to the measured mean and RMS of the signal.

Once these tests are done, modules or fibers which show problems or any bad behavior can be masked or removed from DAQ and from data analysis.

The final fraction of missing fibers and modules has to remain low in order to keep the sensitive area as wide as possible. On the other hand, it is convenient to exclude the bad fibers and modules from the DAQ and/or the reconstruction in order to minimize the possibility of fake hits and noise in the detector.

2.3.4 The latency scan procedure

In order to collect physics data in the best possible conditions it is convenient to ensure that the signal is readout in the moment in which it is at the maximum. In fact the signal shape in peak mode is, ideally, the transfer function in the time domain of a CR-RC circuit:

$$S_{peak}(t) \propto \frac{t}{\tau} e^{-\frac{t}{\tau}} \tag{2.2}$$

where τ is the rise time. The pulse from the APV amplifier lasts about 300 ns, which is large with respect to the time interval of 25 ns between two subsequent bunch crossings.

The APV stores a voltage proportional to the input charge in an internal memory pipeline every 25 ns. The latency of the moment in which the storage of this value happens determines the pipeline cell containing the maximum charge from the traversing particle. During the commissioning (*timing*) runs the synchronization of all the Tracker modules is done in order to make the storage for all of them happening in the same pipeline cell, accomodating the delays introduced by the hardware configuration (fiber lengths, CCU configurations, FEDs, etc.).

2.4 Local reconstruction

The local reconstruction is done transforming the digitized information from the silicon strips into reconstructed hits in the local coordinate of the silicon sensors. The resulting reconstructed hits are input to the various track reconstruction algorithms, which are described in Sec. 2.5.

The raw data coming from the readout electronics of the silicon strip detectors are unpacked and grouped according to the detector modules. After the unpacking step, the raw data are commonly referred to as *digis*, which denotes pedestal-subtracted and zero suppressed ADC counts for individual strips. The digis are associated with a *detector ID*, a unique number which encodes the location of each module in the numerical structure of the CMS tracker. Starting from the digis, the local reconstruction is performed in two stages: cluster reconstruction and hit conversion.

2.4.1 Cluster reconstruction

The cluster reconstruction groups adjacent strips whose associated charge pass a set of thresholds. The thresholds depend on the noise levels characterizing the strips of the cluster. Clusters are reconstructed by searching for a seed strip with a signal-to-noise ratio (S/N) greater than 3. Neighboring strips are attached to the cluster if their signal-to-noise ratio exceeds 2. The total signal size of the cluster, which is the sum of the ADC counts of all associated strips, must exceed five times the quadratic sum of the individual strip noises. The cluster reconstruction algorithm is referred to as *3-Threshold* algorithm and is sketched in Fig. 2.9.

Once the gain calibration is performed, the corrections for the electronic gain can be applied, giving a more appropriate value for the signal. Nevertheless the algorithm is based on the signal-to-noise ratio which is not affected by the gain calibration. For this reason the signal-to-noise ratio is also the most appropriate quantity to study in order to monitor the performance.

2.4.2 Hit conversion

The hit conversion associates every cluster with a hit position and corresponding errors. The hit position is determined from the centroid of the signal heights. The position resolution is parametrized as a quadratic function of the projected track width on the sensor in the plane perpendicular to the strips [66]. Deviations from the ideal geometry ("misalignment") are taken into account by adding an additional uncertainty on the module positions (Alignment Position Error, APE) to the hit errors. The size of the APEs was initially estimated from survey data [67]. Once the alignment with real data can be done and correction constants become available, the APE can be reduced and it can reach values of about 150

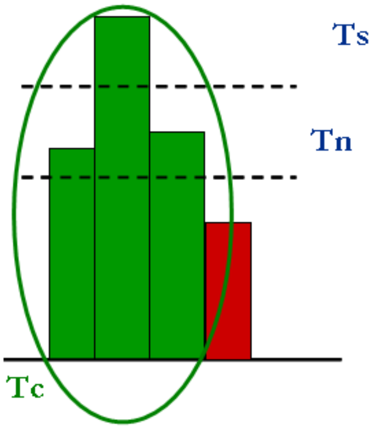


Figure 2.9: *Sketch of the clusterizer 3-thresholds algorithm. The rectangles represent the signal on the strips. The highest green rectangle has passed the seed threshold, therefore the two adjacent strips that pass the neighbor thresholds are attached to the cluster. The red rectangle instead represents the signal on a strip that hasn't passed the neighbor threshold so it is excluded from the cluster. Hence, the sum of the three signals represented in green is calculated, and this is the final signal of the whole cluster.*

μm in the TOB and 500-600 μm in the other subdetectors.

2.5 Track reconstruction

For the LHC collisions two standard algorithms have been developed: "Combinatorial Track Finder" (CTF) and "Road Search" (RS). Both use the hits described in the previous section. The position estimates may depend on the local track angles. In addition, a reconstruction geometry describing the location of the modules and the distribution of passive material and condition information about the status of the different modules are needed.

The tracking algorithms decompose the task of track reconstruction into three stages:

1. seed finding, which provides a selection of initial hits and a first estimate of parameters,
2. pattern recognition, which associates hits to a track, and
3. track fitting, which determines the best estimate of parameters.

The first two items are specific to each of the algorithms while the track fit is always performed by a Kalman filter and smoother. All these software modules use some common services. Material effects - energy loss and multiple Coulomb scattering - are estimated each time a track crosses a detector layer. The amount of material at normal incidence is obtained via the reconstruction geometry.

A third track algorithm has been developed particularly for the cosmics data taking: the Cosmic Track Finder. It is optimized to reconstruct single track cosmic events, but not for the reconstruction of cosmic showers. The stages of the reconstruction are the same as the ones listed for the other two algorithms. The Cosmic Track Finder will be described in detail in next chapter.

2.5.1 Combinatorial Track Finder

The Combinatorial Track Finder is based on the Combinatorial Kalman Filter (CKF), which uses the capacity of the Kalman Filter [68] for simultaneous pattern recognition and track fitting. Starting from an initial estimate of the track parameters the algorithm iterates through the layers of the tracker and builds a combinatorial tree of track candidates.

Seeds are created in the innermost layers of the tracking system. A seed is made out of a hit pair and a loose beamspot constraint or out of a hit triplet. The starting parameters of the trajectory are calculated from a helix passing through the three points. The selected hits must be pointing towards the interaction point and a minimum transverse momentum cut is applied.

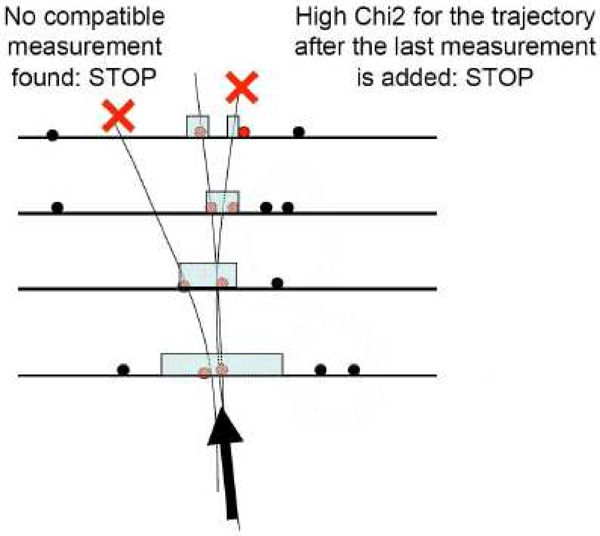


Figure 2.10: *Kalman Filter based pattern recognition.*

The pattern recognition, that is the building of a candidate trajectory by selecting hits out of all the hits in the event, is the computationally most time consuming part of the track reconstruction. It is sketched in Fig. 2.10. From each seed a propagation to the next surface is attempted. Hits are identified in a window whose width is related to the precision of the track parameters. If a hit is found within a window, it is added to the candidate trajectory and the track parameters are updated. If several compatible hits are found, a new candidate is created for each of them. Candidates are sorted according to quality (basically on the χ^2 and the number of hits) and the five best ones are retained for further propagation. As hits are added to the candidate trajectory the knowledge of the track parameters improves and the sizes of the search window decrease. The propagation of a candidate ends if configurable cuts on the number of layers or the number of consecutive layers without a hit are exceeded.

In this contest it is evident how important can be the reduction of the fake clusters in order to optimize the tracking algorithm.

2.5.2 Road Search

For seed finding, the Road Search algorithm uses pairs of hits in seed rings. The set of rings that composes the road will be that consistent with the linear extrapolation between the seed rings in the $r - z$ plane. The seed is composed of a pair of hits in the seed rings within a maximum $\Delta\phi$, which effectively translates to a cut on the minimum transverse momentum of the track. The roads are constrained to point back to the luminous region of the beam, corresponding to roughly $|z| < 15$ cm.

In the first part of the pattern recognition step, an expected trajectory is determined using the two seed hits and the beamspot. The trajectory is extrapolated through the other rings of the road and hits are collected inside a narrow window around the expected trajectory. This collection of hits is referred to as a *cloud*. The cloud should contain all the hits of a track, along with other hits that happen to overlap and lie close to the track. In the second part of pattern recognition, the cloud is turned into a trajectory. A trajectory is first built in low occupancy layers, extrapolating backward. With the trajectory well-defined from the low occupancy layers, hits from the higher occupancy layers are added to the trajectory. The final track will contain at most one hit per detector module, though potentially more than one hit per layer due to detector overlaps.

2.6 The Tracker alignment

The large number of independent silicon sensors (about 15 148, as already stated, plus 1 440 silicon pixel) make the alignment of the CMS strip and pixel trackers a complex and challenging task. The residual alignment uncertainties should not lead to a significant degradation of the intrinsic tracker resolution. Therefore, the required accuracy of the alignment has to be at least equal to, but ideally significantly better than, the ideal track parameter resolution. Certain physics requirements, such as the $W^{\pm}$ boson mass measurement, place even more stringent constraints on the alignment precision. To achieve a desired precision of $\sigma(M_W)$ $<$ 15-20 MeV, the momentum scale has to be known to an accuracy of 0.020-0.025%. This implies the absolute knowledge of detector positions to be known with better than 10 μm uncertainty in the $r - \phi$ plane. This level of accuracy can only be reached with a track-based alignment procedure.

However, track-based alignment cannot be used as the only method for alignment since it relies on tracks and, in turn, pattern recognition for single track reconstruction already requires the position and orientation of tracking devices to be known with an accuracy higher than the mechanical placement precision of the individual tracking subdetectors. Therefore, the alignment has to be performed in 3 steps:

1. measurement of placement and its precision during assembly of tracking devices, e.g., from photogrammetry and detector position survey measurements;
2. measurement of relative positions of subdetectors using the Laser Alignment System (LAS);
3. track-based alignment.

2.6.1 Tracker coordinate systems

CMS uses a right-handed coordinate system, with the origin at the nominal collision point, the x-axis pointing to the center of the LHC, the y-axis pointing upward (perpendicularly to the LHC plane), and the z-axis along the anticlockwise-beam direction. The polar angle (θ) is measured from the positive z-axis and the azimuthal angle (ϕ) is measured from the positive x-axis in the $x - y$ plane, whereas the radius (r) denotes the distance from the z-axis.

In the barrel region, modules are arranged in linear structures parallel to the z-axis, such as ladders in the Barrel Pixel detector (BPIX), strings in the TIB, and rods in the TOB. The endcaps are composed of disks, which in turn contain wedge-shaped structures covering a narrow ϕ region, such as blades in the Forward Pixel detector (FPIX) and petals in the TEC. The BPIX and the TIB are composed of two half barrel structures, separated along the $x = 0$ plane for the BPIX and the $z = 0$ plane for the TIB. A local right-handed coordinate system is defined for each module with the origin at the geometric center of the active area of the module.

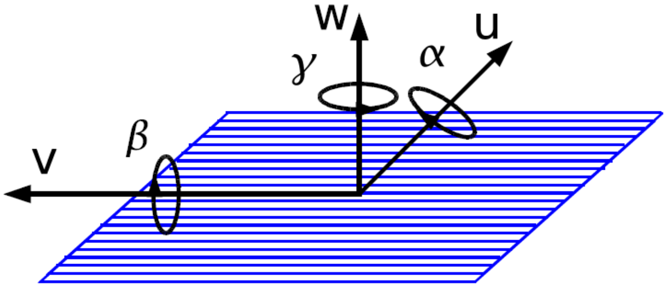


Figure 2.11: *Illustration of the module local coordinates u, v, w and the corresponding rotations α, β, γ for a single-sided strip module.*

As illustrated in Fig. 2.11, the u-axis is defined along the most precisely measured coordinate of the module (typically along the azimuthal direction in the global system), the v-axis orthogonal to the u-axis and in the module plane, pointing away from the readout electronics, and the w-axis normal to the module plane. When double-sided modules are considered as a single entity, the coordinate system is referenced to the $r\phi$ module. For the pixel system, u is always orthogonal to the magnetic field, that is in global $r\phi$ direction in the BPIX and in the radial direction in the FPIX. The v coordinate is perpendicular to u in the sensor plane, that is along global z in the BPIX and at a small angle to the global $r\phi$ direction in the FPIX. The angles α, β, and γ indicate rotations about the u, v, and w axes, respectively.

In addition, local u' and v' coordinates are defined such that they are parallel to u and v, but the direction is always chosen to be in positive ϕ, z, or r directions, irrespective of the orientation of the local coordinate system. For the TID and TEC wedge-shaped sensors, where the topology of the strips is radial, the u'- and v'-axes change direction across the sensor such that v' is always directed along the strips and therefore u' corresponds to the global $r\phi$-coordinate.

2.6.2 General alignment strategy

During the assembly of the CMS Tracker, positions and orientations of the tracker components (silicon sensors as well as support structures) are being measured and stored in a database with the help of Coordinate-Measurement Machines (CMM) and photogrammetry. The former were used for measurements of the active elements, the latter for the alignment of larger structures. Such measurements can be transformed, directly or indirectly, into detector positions, therefore they provide a preliminary correction to the otherwise assumed ideal tracker geometry. If no complete measurement of all silicon detectors is available, because only samples have been taken, the standard deviation of the measurements can be used as an estimate of the placement uncertainty. The error of the track hit position on a corresponding module is enlarged by this value, leading to an improved efficiency during initial track reconstruction.

After assembly, the LAS provides the positions of the tracker subdetectors TIB, TOB, and TEC (not including the pixel detector and the TID) with respect to each other with a precision of about 100 μm.

At this point, an initial track-based alignment procedure can significantly benefit from the good prior knowledge of the positions and orientation of the tracker substructures. In theory, global alignment uncertainties of $\approx 10\ \mu m$ or better can already be achieved with a sample of tracks obtained from only few days of running at low luminosity. In practice, however, the time dependent movements of individual detector components, as well as the systematic uncertainties involved in the track-based alignment algorithm, are the limiting factors for the tracker alignment. A combination of the LAS and the track-based alignment is therefore used to monitor the time-dependence of the tracker alignment due to shutdown operations, dry-out effects, or thermal stress during operation.

Once the tracker is aligned, it defines the reference system relative to which the positions of the calorimeter modules and muon detectors have to be adjusted. In particular, for the optimal reconstruction of muon tracks with high p_T, it is important to combine information from the tracker and the muon detectors.

2.6.3 Track-based alignment

Track-based alignment is absolutely necessary to achieve a sensor precision of the order of a few microns. The goal in track based alignment is to determine the module positions from a large sample of reconstructed charged particle trajectories.

This goal is achieved minimizing a global χ^2 variable:

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_{j}^{tracks} \sum_{i}^{hits} \mathbf{r}_{ij}^T(\mathbf{p}, \mathbf{q}_j) \mathbf{V}_{ij}^{-1} \mathbf{r}_{ij}(\mathbf{p}, \mathbf{q}_j) \qquad (2.3)$$

In this equation, $\mathbf{r}$ is the residual, the difference between the track extrapolation and the hit; $\mathbf{V}$ is the residual covariance matrix; $\mathbf{p}$ and $\mathbf{q}$ are the sensor and track parameters, respectively. The CMS tracker alignment employs two statistical methods: Millepede II [71], a global method using global matrix inversion, and HIP (Hits and Impact Points)[69], a local iterative method.

Millepede II minimizes the χ^2 function in eq. 2.3 by taking into account track and alignment parameters simultaneously. Since only the n alignment parameters $\mathbf{p}$ are of interest, the problem is reduced to the solution of a matrix equation of size n. The value of n is of the order of 10^5 for six degrees of freedom of 16 588 modules. The covariance matrix $\mathbf{V}_{ij}$ is approximated by a diagonal matrix with uncertainties σ_{ij} for uncorrelated hit measurements $\mathbf{m}_{ij}$ of the track j. Given reasonable starting values, $\mathbf{p}_0$ and $\mathbf{q}_{j0}$, the track model prediction $\mathbf{f}_{ij}(\mathbf{p},\mathbf{q}_j)$ can be linearized. Since angular corrections are small, the linearized problem is a good approximation for alignment.

The first step of the global method aligns the highest-level structures (half-barrels, endcaps) with all six degrees of freedom together with all module units, including $r\phi$ and stereo strip modules in a double-sided module, with the most sensitive degrees of freedom each (u, w, γ, and for pixel modules also v).

In the second step all modules (all double-sided or single-sided strip modules with more than 150 hits and all pixel modules with more than 25 hits) are aligned in the TIB in u, w, α, β, γ; in the pixel system in u, v w, γ; and in u, w, γ elsewhere. Compared to the previous step, this recovers the remaining 10% of the modules and allows more degrees of freedom for the TIB, which has larger assembly tolerance, but does not allow independent alignment of the $r\phi$ and stereo modules within double-sided combinations.

The third step is designed to recover lost correlations between the first two steps and has the same configuration as the first step, but the minimum number of hits in the strip modules is increased to 450 with respect to 425 used in the first step.

HIP approximates eq. 2.3 by assuming no track parameter $\mathbf{q}$ dependence and therefore ignores correlations between alignment parameters for different modules in one iteration. The trajectory impact point $\mathbf{f}_{\alpha j}$ is recalculated for each hit, removing the hit under consideration from the track fit. The track parameters and correlations between different modules are taken into account through iterations of the minimization procedure and refitting the tracks with new alignment constants after each iteration. The number of iterations is determined empirically. This approach allows a simplification of the formalism and leads to a solution of the six-dimensional matrix equation for the six alignment parameters of each module in each iteration.

The local iterative algorithm permits the inclusion of survey measurement in the formalism of eq. 2.3. Including the survey constraints, eq. 2.3 becomes

$$\chi^2_m(\mathbf{p}) = \sum_i^{hits} \mathbf{r}_i^T(\mathbf{p})\mathbf{V}_i^{-1}\mathbf{r}_i(\mathbf{p}) + \sum_i^{survey} \mathbf{r}_{*j}^T(\mathbf{p})V_i^{-1}\mathbf{r}_{*j}(\mathbf{p}) \tag{2.4}$$

The additional term is to be minimized independently for each module in a given iteration. In this equation, j runs over all tracker survey hierarchies, which are at the module level in the pixel detector and in

TIB and TID, whilst in TOB and TEC the survey is only done at the level of higher structures.

The track residuals $\mathbf{r}_i(\mathbf{p})$ do not have explicit dependence on track parameters and enter the sum over hits in a given module. The six-dimensional survey residuals $\mathbf{r}_{*j}$ are defined as the difference between the reference and the current sensor position. The survey measurement covariance matrix $\mathbf{V}_*j$ reflects both the survey precision and additional uncertainties due to changes in the detector. These errors can be configured differently for different hierarchy levels and for the degrees of freedom that should be stable, such as the longitudinal direction in a barrel ladder, and those which may change more with time.

The local iterative method uses the full implementation of the Kalman filter track reconstruction algorithm adopted in CMS. It requires a large number of iterations and large computing resources to refit the tracks in each iteration. The global method, instead, allows the determination of alignment parameters, properly accounting for the correlations among them, in a single step. However, the global method does not take into account the effects of material in the tracker and assumes a simple helical trajectory for charged particles. The method also requires a large amount of computer memory and the application of advanced techniques for solving eq. 2.3. Each of the two alignment algorithms can be used to obtain module positions independently and a comparison of results between the two complementary approaches is part of the validation of the tracker alignment procedure.

Chapter 3

The Silicon Strip Tracker commissioning with cosmic rays data

The silicon strip modules, during their full path from the production sites up to their final integration in the tracker and then the insertion in CMS experiment, have been deeply and repeatedly tested to ensure they are in the best possible working condition.

First experience with tracker operation and track reconstruction was gained during summer 2006, when elements of the silicon strip tracker were operated at room temperature in a comprehensive slice test (called *Magnet Test and Cosmic Challenge - MTCC*) involving various CMS subdetectors [65]. The tracker setup was limited and represented by only 1% of the total electronic channels and an active area of 0.75 m^2. Cosmic rays detected in the muon chambers were used to trigger the readout of all CMS subdetectors. The CMS superconducting solenoid provided a magnetic field of up to 4 T. Over 35 million events were recorded and the tracking performance was studied both using a dedicated algorithm for cosmic rays tracking and a general algorithm for track reconstruction in LHC collisions. In addition, tracks reconstructed in the silicon strip tracker were compared with tracks detected by the muon chambers.

In the period from November 2006 to July 2007 the different subsystems of the silicon strip tracker were integrated and commissioned in a large clean room at CERN, the Tracker Integration Facility (TIF). The data taking performed once the whole tracker had been assembled is often referred to as Sector Test [1]. As part of the commissioning, large samples of cosmic ray data were recorded under different running conditions. No magnetic field was present, and the tracker setup consisted of up to 15% of the electronic channels. Over 4.7 million events were taken while operating the detector at five different temperature points. The data were used to verify the reconstruction and calibration algorithms for the LHC collisions modified for cosmic rays tracking.

After this relatively short test on a section of the whole detector, the tracker was moved to its final position, inside the CMS detector, in December 2007. Then the commissioning procedure was repeated in the final environment starting from spring 2008, and the cosmic data taking during CMS Global Runs (runs with the whole CMS detector switched on) started in July 2008 and lasted almost four months. More than 35 million cosmic muons data were collected in the tracker, and many analysis were done to calibrate the detector in its final configuration, with the magnetic field switched off (CRuZeT - Cosmic Run at Zero Tesla) and on (CRAFT - Cosmic Run At Four Tesla). Comparisons with Monte Carlo simulations were also done in order to check that the performance of the tracker was as expected.

In this section I will describe all the analysis done during the two data taking in which I took part: the Sector Test at TIF and the Global Runs CRuZeT/CRAFT in CMS.

For both tests I will give an introduction on the experimental setup and a description of the commissioning procedure, which is fundamental before any physics data can be taken. Then, after an illustration of the local reconstruction and the tracking used for cosmic data, I will briefly describe the Data Quality Monitoring (DQM) system, to the development of which I took part. I will focus on the DQM module I developed for the monitoring of the signal-to-noise ratio of the clusters associated or not to the tracks, and I will show the further analysis I did with the output given by this monitoring module.

Finally, for each data taking I will show the results obtained and the experience gained, in some cases also comparing the results with Monte Carlo simulations.

3.1 The Sector Test at TIF

The goals of this test included commissioning the active sectors of the Tracker under realistic cabling and grounding conditions, establishing the data acquisition, and confirming stable and safe operation under the supervision of dedicated monitoring systems. The objectives were also to develop and validate monitoring tools, to test the connections, to establish stable and safe running at low temperature, and to demonstrate operating procedures. Finally, the acquisition of cosmic ray data allowed the measurement of the detector performance, the understanding of tracking algorithms, and to perform an initial alignment of the active modules.

3.1.1 The experimental setup

The test progressed in an incremental way, beginning with testing part of the subsystems, then proceeding to a test of the barrel systems, and finally incorporating one endcap.

For the Sector Test a fraction of the tracker was selected to be read out so as to maximize the crossing of all the different layers or rings by cosmic rays. The first quadrant in the $z+$ side of the tracker was chosen. Details of Sector Test sub-detectors are shown in Table 3.1. Sensors were biased with voltage of 300 V for TIB, TID and TOB, and 250 V for TEC.

Tracker Sub-detector	Number of Modules	Number of Channels	Fraction of Tracker
TIB	438	282 624	16%
TID	204	141 312	25%
TOB	720	476 460	15%
TEC	800	483 328	13%
Total	2162	1 383 424	15%

Table 3.1: *Composition of the Sector Test with the different subdetectors fractions*

The signal was read out and multiplexed using the APV chips used in Peak mode. Data can be taken in two different modes: Virgin Raw (VR) or Zero Suppressed (ZS). In VR mode, all channels are read out with the full 10 bits ADC resolution. In ZS mode, the FED applies pedestal subtraction, common mode rejection and a fast clustering algorithm, using signal height with reduced 8 bits resolution: only channels forming a cluster are output. Almost all cosmic ray data in this phase were taken in Virgin Raw mode since low cosmic trigger rate did not require data reduction, and VR running allowed offline optimization of thresholds (in LHC collisions it will be compulsory to operate in ZS mode).

The cooling plant used for the Sector Test was simpler than the final system and its cooling power limited. A minimum operating temperature of -10°C was obtained, compared to the -25°C in the final system. The temperatures measured at the cooling tubes proved to be very stable with variations of less than 0.1°C. A test was possible at a temperature of -15°C but only by limiting power to half of the Sector Test modules.

Dry air was supplied inside to maintain a sufficiently low dew point to avoid condensation; it flowed in a semihermetic tent that contained the tracker.

The safety of the tracker was constantly monitored through the Tracker Safety System (TSS) and the Tracker Control System (TCS). These two systems constantly monitored the temperatures, the voltages, cooling plant failures, and other alarm conditions. All these systems were prototypes of the ones that will be used during collision data taking that have also been used during Global Runs.

3.1.2 Cosmic Muon Trigger

The cosmic trigger configuration was designed to allow studies of tracking performance and detector alignment. The trigger design was constrained by space above and below the tracker; in particular clearance below the tracker allowed only 5 cm lead bricks for filtering low momentum tracks. Six scintillators (T1, T2, T3, T4, T5, T6) were placed above the tracker, in a fixed position; below the tracker there was initially only one scintillator (B0) mounted on a movable support structure; later a further set of four scintillators was added (B1, B2, B3, B4) to increase the trigger acceptance.

The upper scintillator signals were synchronized by simple NIM logic using cosmic rays and put in a logical-OR to obtain a Top-scintillator signal. This was put in coincidence with the lower scintillator signal. In the same way, from synchronized logical-OR signals of scintillators B1, B2, B3, B4 a Bottom-scintillator signal was obtained and was put in coincidence with the Top-scintillator signal.

Three main trigger configurations have been used at the Sector Test:

- (TA) = (Top-Scintillator & B0) vertical position;
- (TB) = (Top-Scintillator & B0) slanted position;
- (TC) = (TA) + (Top-Scintillator & Bottom-Scintillator) slanted position.

The schematic representations of these three configurations are shown in Fig. 3.1. The rate of spurious coincidences relative to the true level was of order 10^{-9}. The trigger rates achieved were: 3.5 Hz (TA), 1.5 Hz (TB), and 6.5 Hz (TC). Since the DAQ rate was limited to about 3 Hz by the FED readout via VME, a trigger veto was implemented to keep the rate under this level.

3.1.3 Data sets and reconstruction

Depending on the different trigger configurations and operating temperatures, different data sets have been defined, as they are summarized in Table 3.2.

The APV parameters were studied for TIB and TOB detector modules only. For the TID and the TEC (where a large number of variants exist) the APV parameters have not been optimized for each module type: TID used the TIB parameters while the TEC used parameters very close to the TOB ones.

Each run was checked using online and offline quality monitoring tools (see Sec. 3.1.4). If a run did not meet the quality requirements or if a configuration or hardware problem was discovered, it was flagged as bad and excluded from the offline analysis.

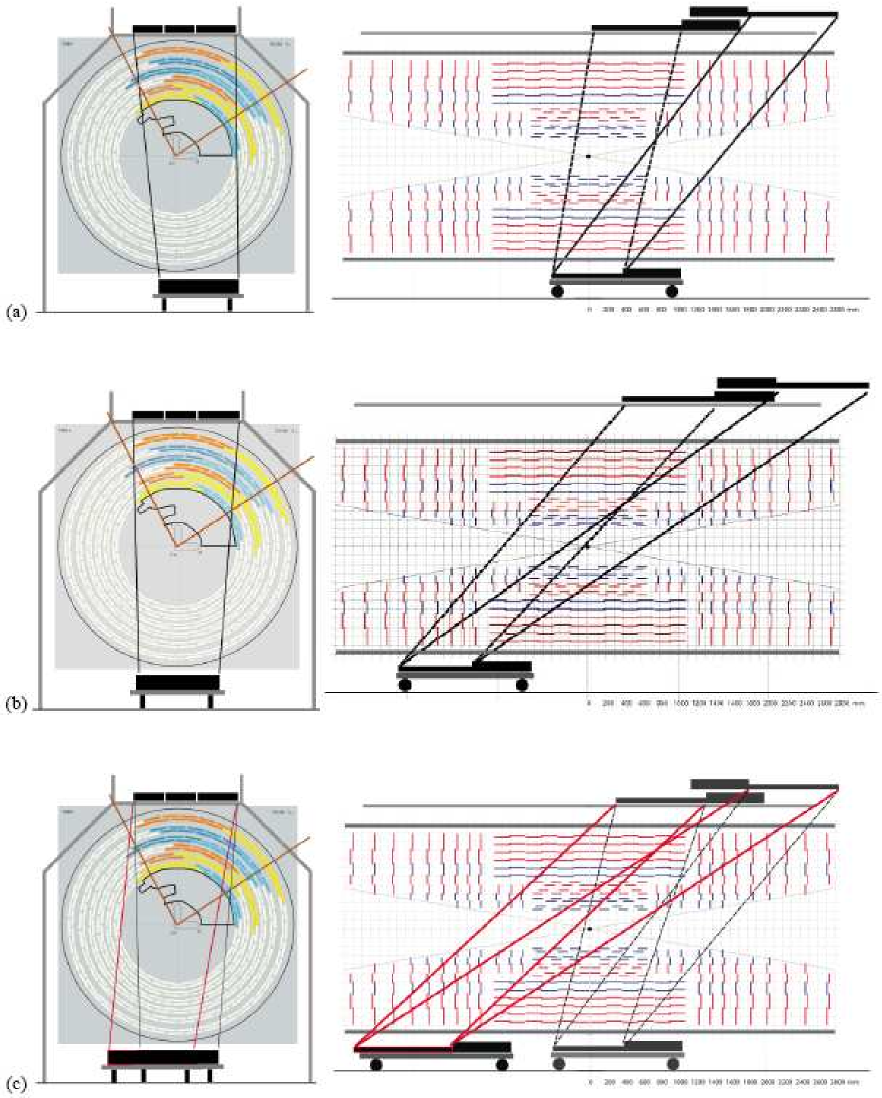


Figure 3.1: *Layout of the various trigger scintillator configurations used during the cosmic data taking at the TIF (in chronological order): (a) configuration TA; (b) configuration TB; (c) configuration TC. The xy view is shown on the left side, the rz view is shown on the right. The straight lines connecting the active areas of the top and bottom scintillation counters indicate the acceptance region.*

The complete data set was processed several times using different software releases to integrate improvements in track reconstruction.

Comparisons have also been done with Monte Carlo simulated samples. The simulation of a realistic cosmic muon spectrum relies on dedicated parametrizations of the energy dependence and incident angle, also accounting for the correlations between the two quantities. For compar-

Run	Trigger	Operating T (°C)	Total Events	Good Events
6203-6930	A	15	703 996	665 409
7277-7296	A	15	191 154	189 925
7635-8055	B	15	193 337	177 768
9255-9341	C	15	132 311	129 378
10145-10684	C	10	992 997	534 759
10848-11274	C	-1	893 474	886 801
11316-11915	C	-10	923 571	902 881
12045-12585	C	-15	656 923	655 301
12599-12656	C	15	112 139	112 134

Table 3.2: *Sector Test data sets. TEC has been tested only starting from run 7277.*

isons with TIF data, cosmic muons have been generated starting from an ideal cylinder coincident with the CMS surface using the CMSCGEN generator [77]. The original parametrization of the spectrum is based on a polynomial fitted to data above an energy of 10 GeV and cannot be used at much lower energies. In order to cover the range from 200 MeV to 2 GeV the cosmic muon spectrum is produced assuming a simple energy dependence in CMSCGEN and reweighted using the CAPRICE [78] energy spectrum. The configuration parameters used for the generation process are listed in Table 3.3. Since the tracker is on the surface without any iron shielding from muon stations outside, the energy loss scale factor was set equal to zero.

Parameter	Value
Min(E)	200MeV
Max(E)	10 TeV
Min(θ)	0°
Max(θ)	88°
Min(ϕ)	0°
Max(ϕ)	360°
Min(T_0)	-12.5 ns
Max(T_0)	12.5 ns
Energy Loss Scale Factor	0.0

Table 3.3: *Parameter settings used for the generation of cosmic muons (θ is the angle between the muon direction and the y-axis in the CMS reference system and T_0 is the allowed muon time of flight).*

Before the detector and electronics response are simulated, a special filter is applied reproducing the trigger setup. The scintillation counters are modeled as virtual 1×1 m^2 surfaces, and the muon trajectories are extrapolated to the outside of the tracker where the intersection points with the scintillator surface are calculated. Using the intersection points, the trigger logic is applied and the cosmic muon event is either retained or discarded. Simulated events are generated separately for all three trigger layouts.

3.1.3.1 Local Reconstruction

Event reconstruction and selection, data quality monitoring and analysis of the Tracker Sector Test at the TIF were performed within the CMS software framework, known as CMSSW [79]. Hit and track reconstruction was performed offline taking the raw data as input. A reconstruction job is composed of a series of applications executed for each event in the order specified by a configuration file. A fundamental part of the processing is the availability of non-event data such as cable map, pedestals, alignment constants and calibration information.

The first step consists in mapping ADC counts for individual strips as they are coming from the FED output, into objects that are uniquely assigned to a specific detector module exploiting the cabling map information stored in the configuration database. In the case of data collected without performing online zero-suppression (VR), the necessary pedestal information must be acquired from the database. At this point, the input data files, whether real data or Monte Carlo simulated events, contain the same information and can be further processed with identical code.

The three-threshold algorithm described in Chapter 2 has been used to form the clusters.

3.1.3.2 Track reconstruction

The tracking was performed using the two tracking algorithms (CTF and RS) described in Chapter 2, modified in order to match the reconstruction of cosmic tracks, plus a third one developed especially for the cosmic tracks: the Cosmic Track Finder.

The CTF for cosmic rays is identical to the one designed for proton-proton collisions except for the seed finding stage, which has been adapted for the reconstruction of cosmic tracks. In particular, no vertex constraint can be applied, since the cosmic rays do not necessarily cross the tracker pointing towards the interaction point, and seeds should be created also in the outer layers, because these layers have higher acceptance for cosmic tracks. For the reconstruction of TIF data, hit triplets are used in the inner or outer parts of the barrel and hit pairs in the endcaps. Hit triplets are checked for compatibility with a straight line: the radius of the circumference passing through the three hits has to exceed 5 m. The different combinations of layers used for seeding are summarized in Table 3.4.

In the RS algorithm, specific roads for cosmic track reconstruction were generated, where the constraint on the extrapolation of the roads was loosened to include any pair of seed rings within the acceptance of

Seeds	Layers
Inner barrel triplets	TIB1+TIB2+TIB3
Outer barrel triplets	TOB4+TOB5+TOB6, TOB3+TOB5+TOB6, TOB3+TOB4+TOB5, TOB3+TOB4+TOB6, TOB2+TOB4+TOB5, TOB2+TOB4+TOB6
Endcap pairs	Any pair of adjacent TEC wheels

Table 3.4: *Combination of layers used for seed finding in the CTF algorithm*

the readout detector. An overview of the inner and outer seed rings for the TIF geometry is shown in Fig. 3.2.

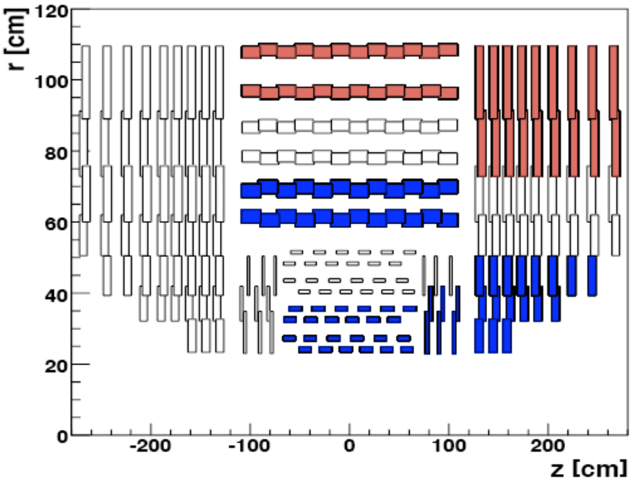


Figure 3.2: *Overview of the ring structure of the CMS tracker. The rectangles in the plot represent the rings as defined in the Road Search algorithm. The inner seed rings are shown in blue, the outer seed rings are shown in red. The seeding is asymmetric since modules located at -z were not part of the cosmic data taking.*

The choice to use the inner layers of both TIB and TOB as inner seed rings (in addition to the inner rings of TID and TEC) was made to avoid any geometrical acceptance loss and to have one seed ring structure which fits different possible sub-detector readout combinations.

The Cosmic Track Finder is instead designed in order to meet the specific task of reconstructing single tracks without imposing a region of origin, but assuming a preferred direction. For cosmic track reconstruction in TIF data, a dedicated seeding is used. The total number of hits in the whole tracking volume is expected to be several orders of magnitude lower than in proton-proton collisions. Hence, all the hit pairs which are geometrically compatible are considered as potential seeds. The definition of compatibility is the following:

- The seed is built from a pair of hits lying in different layers (barrel) or wheels (endcaps). In the barrel, all the hit pair combinations of either the three outermost TOB layers or of the three innermost TIB layers are considered. In the endcap, hit pair combinations are accepted if the hits are separated by at most one intermediate wheel.
- The distance between the modules along the z-axis is less than 30 cm. This cut is not applied for seeds in TEC layers.
- The distance between the hits in global x is less than two times the distance in global y. This requirement is motivated by the small angles cosmic tracks are expected to have with respect to the vertical direction. Again, this cut is not applied to seeds in TEC layers.

All seeds that fulfill the previous selection criteria are considered in the pattern recognition. For each seed, preliminary track parameters are calculated basing on the line connecting the two hits. The hit-selecting algorithm is simplified with respect to the Combinatorial Kalman Filter. A seed, which comes from the previous step, can be at the top or bottom of the instrumented region of the tracker. If it is at the top (bottom), all the hits with a global y coordinate lower (higher) than the hit of the seed are sorted in decreasing (increasing) order with respect to the global y axis. A very simple procedure establishes if the hit can be selected or not:

1. The trajectory is propagated to the surface of the module which provided the hit. The uncertainty from multiple scattering is considered when the track is propagated.
2. The compatibility of the hit with the propagated trajectory is evaluated using a χ^2 estimator. For this analysis a cut at $\chi^2 < 40$ was chosen.
3. If the hit is compatible, the trajectory is updated with the hit.

A trajectory candidate is retained if it contains at least 4 hits. The final selection is only done after the full track fit. The fitting procedure is the same as for all the other CMS tracking algorithms and is based on the Kalman Filter. At the end of this phase several trajectories are still valid, but only one is retained since only one track per event is expected. The best trajectory is selected on the basis of the criteria below, listed in order of precedence:

- the highest number of layers with hits in the trajectory,
- the highest number of hits in the trajectory,

- the lowest value of the χ^2/d.o.f..

The main algorithm used in the analysis I will show is anyway the Combinatorial Track Finder. The other two algorithms have been used to cross check the results.

3.1.4 Data Quality Monitoring

The Data Quality Monitoring (DQM) system for the Tracker is designed to ensure that good quality data are recorded and detector problems are spotted very early on. The system is based on the "Physics and Data Quality Monitoring" framework [80] of CMS. The task is fulfilled in three steps:

- histograms, called Monitoring Elements (MEs), are defined and filled with relevant event information by the "Producer" (DQM Source) application;
- a "Consumer" (DQM Client) application accesses the MEs, performs further analysis and generates alarms;
- the "Graphical User Interface" (GUI) provides tools for the visualization of the MEs.

The MEs are defined in the DQM source at various levels of data reconstruction chain. Starting from the level of pedestal, noise, digitization, cluster reconstruction and finally track-related properties are defined and filled in various MEs. The DQM Client accesses low-level MEs, performs further analysis on them and creates summary MEs, performs quality tests comparing MEs with references to generate alarms. The summary plots are important as it would be too time consuming to check each of the huge set of MEs for the more than fifteen thousand tracker modules. The information from detector level MEs is accessed and summarized in MEs at higher levels following the geometrical structure. Similarly the detector level MEs are compared with reference MEs or parameters and "Ok", "Warning", and "Error" alarms are generated, as appropriate, based on the comparisons. The DQM system was operational with full functionality during TIF data taking. Events were accessed from the on-line data and full reconstruction was performed on an event by event basis and provided to the DQM sources. The system was stable during operations and the tracker DQM GUI, which is a web based application, was accessible from CERN as well as from other institutes involved in the Tracker activities.

The monitored quantities during the TIF data taking were, mainly:

- pedestals and digis;

- cluster properties: signal, noise, signal-to-noise, width;
- track properties: χ^2, number of hits, hit residuals.

Later on, I developed a DQM module to monitor the clusters properties not only in general but also distinguishing them when they are real clusters associated to a track, or noise clusters not associated to any track. This module became available and functional during the Global Runs. During TIF I performed this kind of analysis offline, developing an analysis program to study the properties of clusters, hits and tracks. The results of these studies are illustrated in Sec. 3.3.1.

3.2 Commissioning and calibration at TIF

In Chapter 2.3 a description is given of the calibration procedure to be executed before starting a physics data taking in order to keep all the possible electronic sources of noise or instability under control. This procedure is applied during the commissioning runs, also referred to as timing runs. During TIF data taking, the commissioning runs were repeated at intervals to get more statistics on the stability of the system.

Here the results of the commissioning runs are reported.

3.2.1 Electronic gain measurements

Tick-mark height distributions obtained after the commissioning procedures are shown in Fig. 3.3 for two different temperatures. The different colors represent the components with different AOH gain settings (called 0, 1, 2, 3 from lower to higher gain): the difference between the left and right distributions is mainly due to the strong dependence on the temperature of the AOH gain. The effect is manifest by the increased number of AOH with gain equal 0 at lower temperature. This implies that the average values of the tick-mark height distribution changes when varying the temperature.

Even after the commissioning procedure the tick-mark height distributions still indicate about a 10% spread in electrical gain. This variation is consistent with the coarse precision of the AOH laser gain settings, as already stated in the previous Chapter. Therefore the correction factor of equation 2.1:

$$C_{corr} = \frac{640}{TickMark}$$

should be applied in all the analyses of signal and noise.

One limitation in the use of this equation is due to the different module operating voltages for different layers. It was known that the tick-mark amplitude is linearly proportional to the 2.5 V operating voltage. The signal at the APV amplifier output is not much affected by changes in the

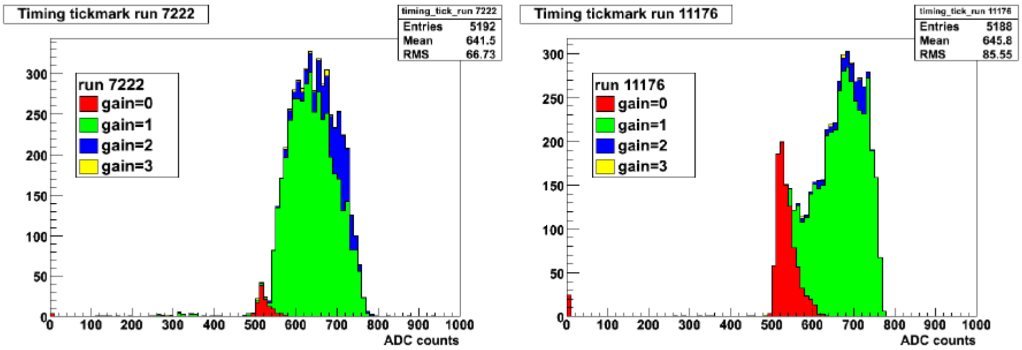


Figure 3.3: *APV tick-mark height distribution after commissioning procedure, run 7096 at T = 14° C (left) and run 11176 at T = -10° C (right)*

supply voltage. Therefore to make a precise estimate of the electronic gain it is necessary either to equalize all the operating voltages or to correct the tick-mark value for the difference compared to 2.5 V. In the following results the correction was not applied, therefore the electronic gain can be considered having a systematic variation of about 5%, which can affect direct comparisons between different layers.

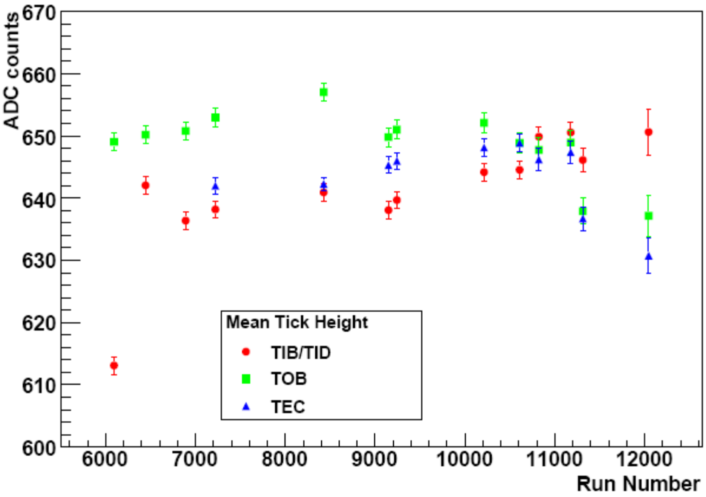


Figure 3.4: *Tick-mark height distribution versus run number, during TIF data taking. Error bars represent the statistical error on the average value.*

Fig. 3.4 shows average tick-mark heights for individual sub-detectors as a function of run number (with the help of Table 3.2 the period at different temperatures can be identified). At least one timing run was taken as part of the commissioning procedure whenever the coolant temperature changed. During the period at T = 14/15 °C (runs 6203-9341), in particular, there were several timing runs, which provide information on the stability of the system.

TOB shows, in 7 runs, variations of ±0.6%. The TIB/TID shows a large discrepancy in the first run (6094), where results are lower by 5%, but this is due to the fact that the modules were operating at low electrical gain and therefore that was corrected by a very high AOH gain: after removing this run the TIB shows, in 6 runs, variations of ±0.6%. The TEC shows, in 4 runs, ±0.4% variations.

3.2.2 Noise performance studies

The noise has been first of all renormalized for each pair of adjacent APV, belonging to the same AOH laser, in order to take into account the different electronic gains, by multiplying the noise for the correction factor specified by equation 2.1. Modules with known problems that affect the noise distribution calculation have been removed from the analysis at least for the period when the problem was present.

For each strip noise distribution, a fit to a Gaussian has been performed and results for some layers are shown in Fig. 3.5.

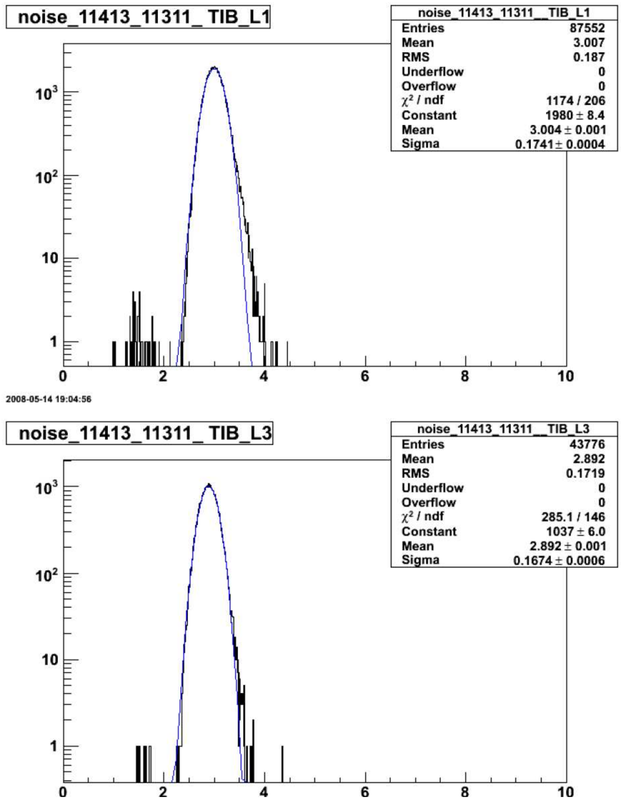


Figure 3.5: *TIB L1 (top left) and L3 (bottom left) noise profile at* T = *-10°. Gaussian fit is shown. On the* x *axis the noise is reported in ADC counts.*

For most of the layers the noise distribution is very well represented by a Gaussian, and fitted values and sigmas are almost identical for identical layers, showing that extra noise sources do not affect the detector performance significantly. The only relevant non-Gaussian tails at high values are visible for TOB layers 2, 3 and 4, an example showed in Fig. 3.6. They are mainly due to channels close to APV edges and only for modules at specific positions within a rod, closer to a clock distribution board and a power cable.

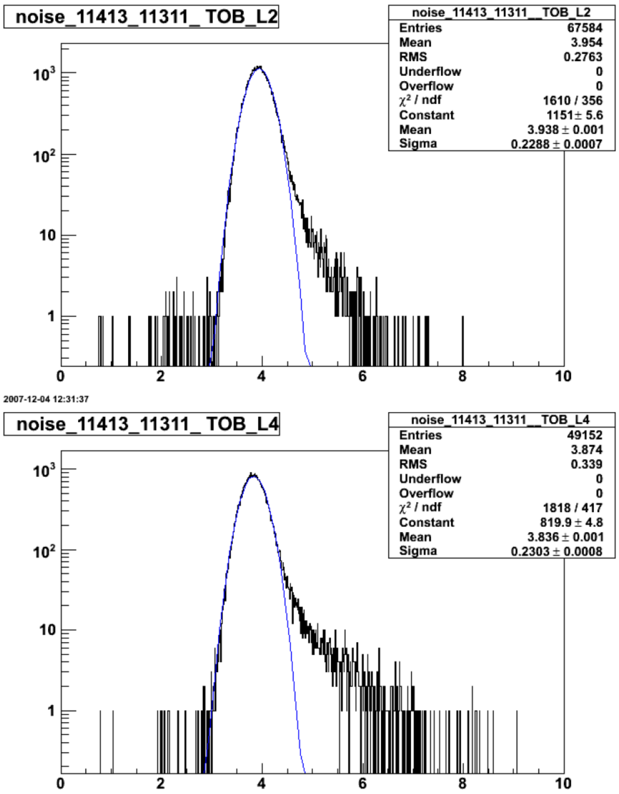


Figure 3.6: *TOB L2 (top figure) and L4 (bottom figure) profile for T = 10°C with evident tail. Gaussian fit is shown. On the x axis the noise is reported in ADC counts.*

This noise pickup does not affect the TOB performance, given its high signal to noise ratio, as it will be shown in the following section. Nevertheless, during the Sector Test many possible grounding and filtering schemes were investigated in order to minimize this extra noise. A solution which was found to be very effective consisted in grounding the TOB power cable shields at the patch panel close to the tracker. This grounding implementation was possible on the Sector Test setup only for a fraction of the TOB, therefore the tails remained for most of the module rods. This grounding scheme has subsequently been implemented during installation of the tracker inside CMS.

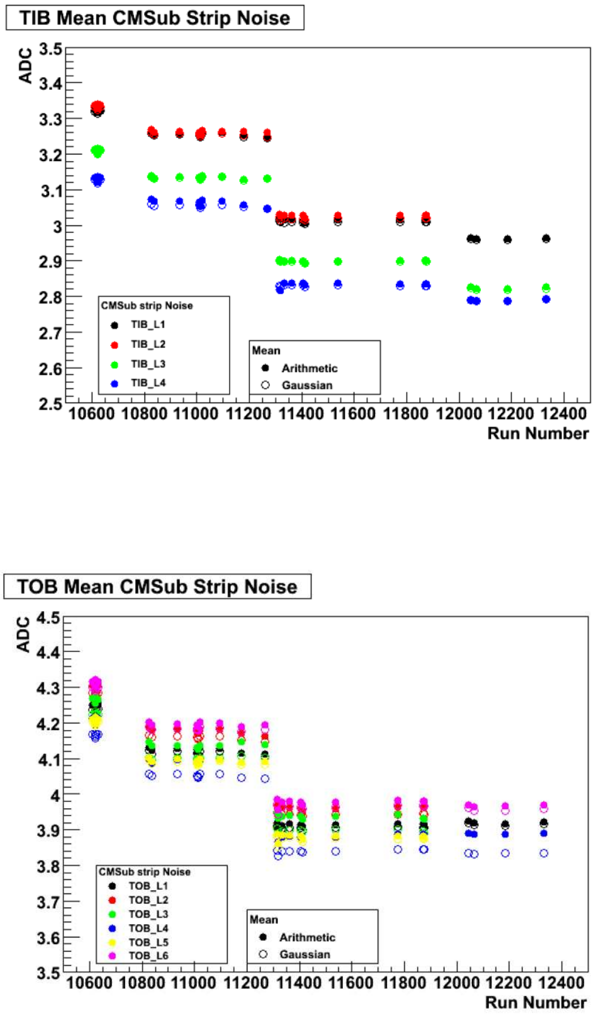


Figure 3.7: *Noise vs run number for TIB (top figure) and for TOB (bottom figure). Four periods at different temperatures are visible: 10 °C, 0 °C, -10 °C and -15 °C.*

Stability of the noise performance was studied by taking pedestal and noise runs at different times when the tracker was running in stable conditions with fixed electronics configuration settings. Results are shown in Fig. 3.7 displaying noise of all layers/rings versus the run number for TIB and TOB: the steps in the trend represent the different temperatures considered, 10°C, 0°C, -10°C and -15°C. The data average and mean value of a Gaussian fit are displayed as solid and open symbols respectively. For constant coolant temperatures the noise is stable to better than ±0.5%. Most importantly, the noise decreases with the decreasing temperature as expected by the laboratory studies made on the APV performance.

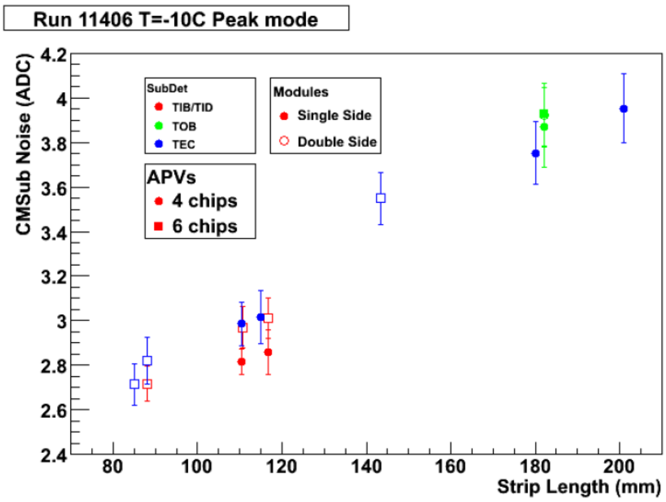


Figure 3.8: *Noise vs strip length for T = -10°C.*

The mean strip noise obtained from a Gaussian fit for each different type of module has been correlated with the module strip length. Results are shown in Fig. 3.8 for a run taken at -10°C, but similar results are obtained at the other temperatures. Error bars represent the spread of the mean module noise for each module type. As expected the behavior is well represented by a linear behavior, within statistical fluctuation, for modules from the same layer, although the average values show differences for the same strip length but different module geometries. This can be explained by the difference in supply voltage and by the different APV settings used for different sub-detectors.

3.2.3 Detector quality analysis

A first identification of bad components is done during timing runs: the dead components are identified looking for low tick-mark heights. Possible dead components correspond to the broken fibers whose channels showed problems during the timing runs. The number of missing fibers in the tracker was found to be at the 0.1% level during Sector Test.

Remaining bad components can be identified complementary to the noise measurement, studying the normality of the noise distribution in order to remove strips with non-Gaussian noise behavior that may lead to unwanted higher occupancy after zero suppression. The procedure described in Sec. 2.3.2 has been applied to isolate noisy or bad strips. Isolated bad channels were identified having higher than five sigmas (“noisy”), or lower than five sigmas (“dead”), noise compared to the average noise per module, and this was done for each module geometry of the tracker.

The number of dead channels is almost constant among several runs for all sub-detectors, showing that the identification of dead channels is clear and stable.

The noisy components are instead subjected to fluctuations, in particular for TOB and TID. The fraction of dead (noisy) strips is 0.05% (0.04%) for TIB, 0.04% (0.15%) for TID, 0.04% (0.3%) for TOB, and 0.08% (0.02%) for TEC.

Since the analysis of defective strips is made on a per run basis, it is important to understand the number of runs in which a strip was identified to be bad. Dead strips identification is stable: the majority of dead strips (70%) were flagged in all runs. About 30% of the classified dead strips appear only in a single run which likely had a timing issue or other unusual problems. On the contrary, only a small fraction of the noisy strips were noisy throughout the Sector Test. In most cases anomalous noise persists for one or two runs at the most. These runs may reflect special operating conditions or non-optimal system configuration. Finally, a comparison of the identified faulty channels with data from the Tracker construction database shows that 90% of the dead channels and 40% of the noisy channels had been flagged as such by the end of the construction period.

3.2.4 Latency scan procedure

During the Sector Test, almost after every significant change in the APV settings or temperature, latency scans were done in order to determine the new latency parameters. Typically 50 events were sufficient for each latency step during the scans. About ten steps in latency were performed, for a total of about 500 triggers on average.

Fig. 3.9 shows the results from one of the latency scans. The latency which maximizes the signal peak for each single subdetector is obtained from fits to these plots. The optimal latencies for the four subdetectors differ from each other for small values, and this is due to the different lengths of the readout fibers from the front-end hybrid to the FED. No tuning of the pulse shape was performed, therefore deviations from an ideal RC-CR shape are present. To take into account the non-nominal behavior of the pulse shape a smearing term has been introduced into the fit. This smearing term obtained from the fit results to be larger for TEC, since the pulse shape of individual APVs was not optimized individually for the many different TEC modules geometries.

The estimated difference between the optimal latency and the latency used in the various phases of the TIF Sector Test is shown in Fig. 3.10. During the many different operating periods of the Sector Test, the commissioning process generally resulted in timing well within a 25 ns window. In one week of running, an incorrect latency value was determined by mistake, which determined an error in timing of about 30 ns. A

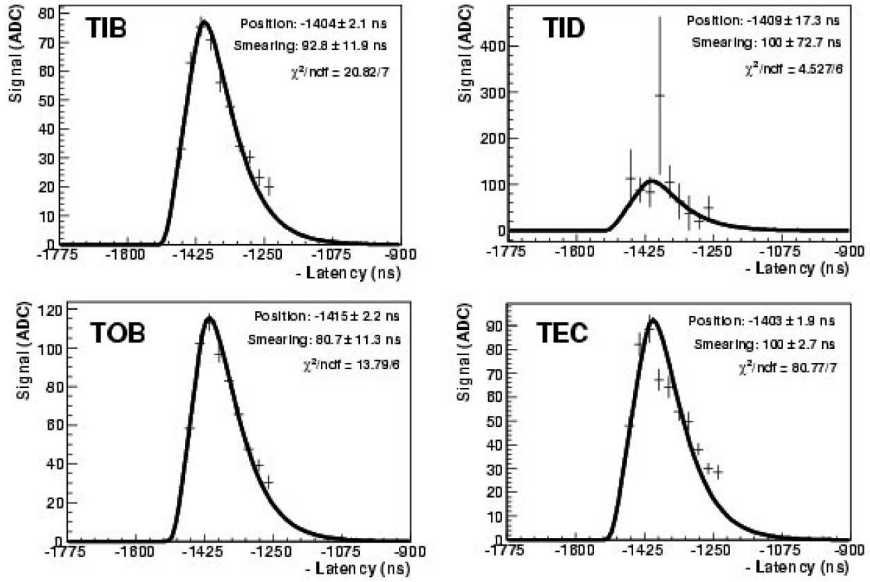


Figure 3.9: *Fit of the signal as a function of the latency, for each of the four sub-detectors. Data are from run 10947 and were used to set the latency for the following period.*

systematic relative difference of about 10 ns between the optimal timing for TOB with respect to TIB/TID and TEC is also visible. The statistical error of these measurements has been estimated to be about ±1 ns.

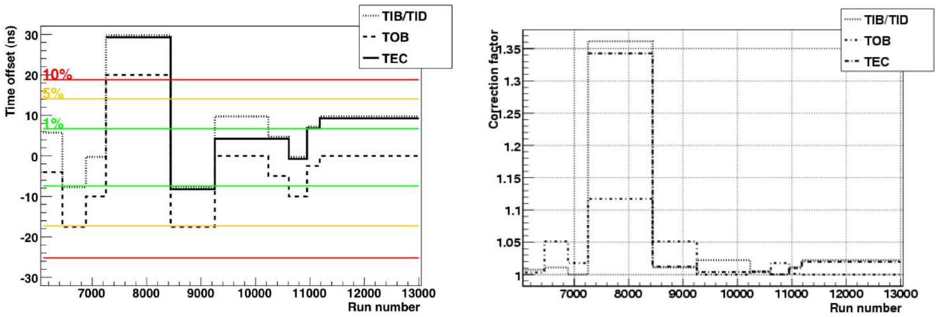


Figure 3.10: *Offset from the on-peak timing (left) and signal correction factor (right) as function of the run number for the three sub-detectors TIB/TID, TOB and TEC.*

The non optimal sampling determines the loss in the signal performance of the various sub-detectors. On the right side of Fig. 3.10, the horizontal lines mark the offset corresponding to a loss of 1%, 5% and 10% respectively. The correction factor to compensate for the loss of signal is plotted for each run on the left side of Fig. 3.10. The anomalous week where a wrong latency was chosen, resulted in large correction factors: 10% for TOB and 35% for TIB/TID and TEC. In the other periods, the correction is less than 5% and this sets a scale for the expected accuracy of the absolute calibration. This means that comparing results between different sub-detectors or for the same sub-detector at different temperature conditions, this level of uncertainty is larger than, for example, the contribution from time-of-flight differences.

3.3 Cosmic rays analysis at TIF

The signal performance of the Tracker is fundamental; it depends on several factors: charge collection in the silicon sensors, performance of the APV with optimized parameters, performance of the full electronic chain.

As the tracking efficiency relies on a sufficiently large signal-to-noise ratio (S/N) in the tracker modules, it is important to measure the S/N on a layer by layer basis. Moreover, S/N analysis is independent on the gain calibration. On the other hand, analysis of the signal size allows a direct comparison of the performance of the different layers as a mean of determining the absolute gain calibration and measurements of temperature dependence.

The energy (S_{tot}) deposited in the tracker modules can be parametrized as:

$$S_{tot} = \frac{dE}{dx} tK; K = \frac{1}{\cos(\theta_{3D})} \tag{3.1}$$

where θ_{3D} is the angle of incidence of the track with respect to the silicon detector normal, as shown in Fig. 3.11, and t is the silicon thickness.

In the following analyses signal values are normalized to the thickness of silicon that the particle has actually traversed.

$$S_{ren} = \frac{S_{tot}}{K} = \frac{dE}{dx} t. \tag{3.2}$$

Fig. 3.11 illustrates the definitions of two of the three angles which will be used in the following analysis. The angle made by the track in a plane orthogonal to the sensor surface and whose X axis is transverse to the strip direction is denominated XZ angle. This is directly correlated with the cluster size. The YZ angle is instead defined in a plane perpendicular to the surface and oriented along the strip direction.

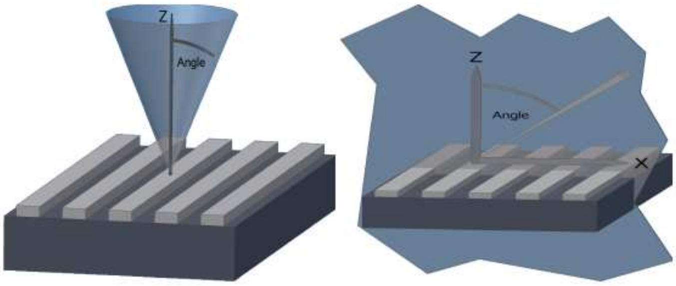


Figure 3.11: *Angle definitions:* θ_{3D} *on the left;* θ_{XZ} *on the right.*

3.3.1 The signal-to-noise ratio

The signal-to-noise renormalized to the sensor thickness (S_{ren}/N) is an ideal parameter for measuring the stability of the tracker during a data taking, since it should be largely independent of gain corrections, and therefore allows an accurate comparison of results from modules in the same layer or at different temperatures from run to run.

S_{ren} is the signal as defined in eq. 3.2. According to the theory of the release of charge in thin detectors, the charge released by a particle in a layer of silicon strips should be distributed accordingly to a Landau shape.

N is the cluster noise, defined as $N = \sqrt{\sum_i N_i^2/n_{strips}}$ where N_i is the noise of the i-th strip of the cluster and n_{strips} is the number of the strips of the cluster. The noise of all the strips of a module has in general a Gaussian distribution, as shown in Sec. 3.2.2. It should be emphasized that if the noise is constant for all the strips in the cluster, then the cluster noise is independent of the cluster size and equal to the strip noise. N_i is determined in each pedestal run and written to the offline database. It is not remeasured continuously during the cosmic data taking.

Another issue that should be put in evidence is that the clusters we are referring to in eq. 3.2 should be the real signal clusters, i.e. the clusters which are actually associated to a track left by a particle passing through the tracker. This distinction is important because there is a given amount of cases in which fake clusters can be formed, mainly for the presence of noisy strips that produce a signal over the zero suppression threshold. It can happen that some of these strips are adjacent, and that they pass the thresholds for the clusterization. In this case, there will be a given amount of clusters that are not associated to any track.

Usually it is easy to isolate the “fake” clusters from the real ones, as they have a very low S_{ren}/N (below ∼10-15) with respect to the signal clusters ($S_{ren}/N > 15$) associated to particles passing through silicon and releasing a large amount of charge. Hence the signal-to-noise ratio of fake clusters will mainly be visible as a spike at the low edge of the Landau distribution of the clusters S/N.

In order to measure the signal-to-noise ratio during TIF data taking I developed a tool which permitted to perform several and specialized studies on the clusters, distinguishing the cases in which they were real signal clusters or fake noise clusters.

This analysis program, developed in the environment of the CMS framework (CMSSW), reads separately the hits associated to tracks in the event and the hits not associated to any track. Then from both this kind of hits it extracts the clusters and all the information related: the charge released by the ionizing particle, the signal renormalized to the thickness, the noise, the ratio between them, the width, the position. Each of these quantities is then written into histograms separately for the different subsystems of the tracker, which differ from each other for dimension, geometry and thickness. These histograms are then saved in an output file, that can be then used for further analysis. For instance, is then possible to perform fits on the signal or signal-to-noise shape to extract the peak value and monitor its stability in time.

The S_{ren}/N distributions have been fitted with a function obtained from the convolution of a Landau function and a Gaussian distribution. The reason of this choice is that the presence of the Gaussian noise smears the Landau shape of the signal.

In order to obtain a high purity signal, only those hits that are associated with a reconstructed track have been used. Nevertheless, some cluster with very low S_{ren}/N, probably noise one, can still happen to be on a track trajectory, therefore ending up in this distribution. This fact can cause the presence of a residual low charge peak also in the S_{ren}/N distribution even if only the hits associated to tracks are considered. In order to avoid biasing the fit considering this region, some tests have been made to optimize the fit range. Finally it was decided to perform two subsequent fits. In the first one, the range was set from about 6% of the mean value up to the mean plus three times the R.M.S. of the distribution. In the second iteration of the fit, once a first estimate of the peak of the Landau, its FWHM, and the sigma of the Gaussian convoluted is available, the range has been set from a minimum value equal to the 10% of the peak to a maximum equal to the peak position plus three times the FWHM.

The fits were performed separately for the different layers of the barrel and the different rings of the endcaps, in order not to mix different geometries and thicknesses, which would cause a deformation of the Landau distribution.

The CTF track algorithm was used and only tracks with $\chi^2/d.o.f. < 30$ were considered. Only events with low track multiplicity (less than three) and with a low hit multiplicity (less than 100) were considered.

In Fig. 3.12 the fit of the S_{ren}/N at T = -10°C is reported for TIB

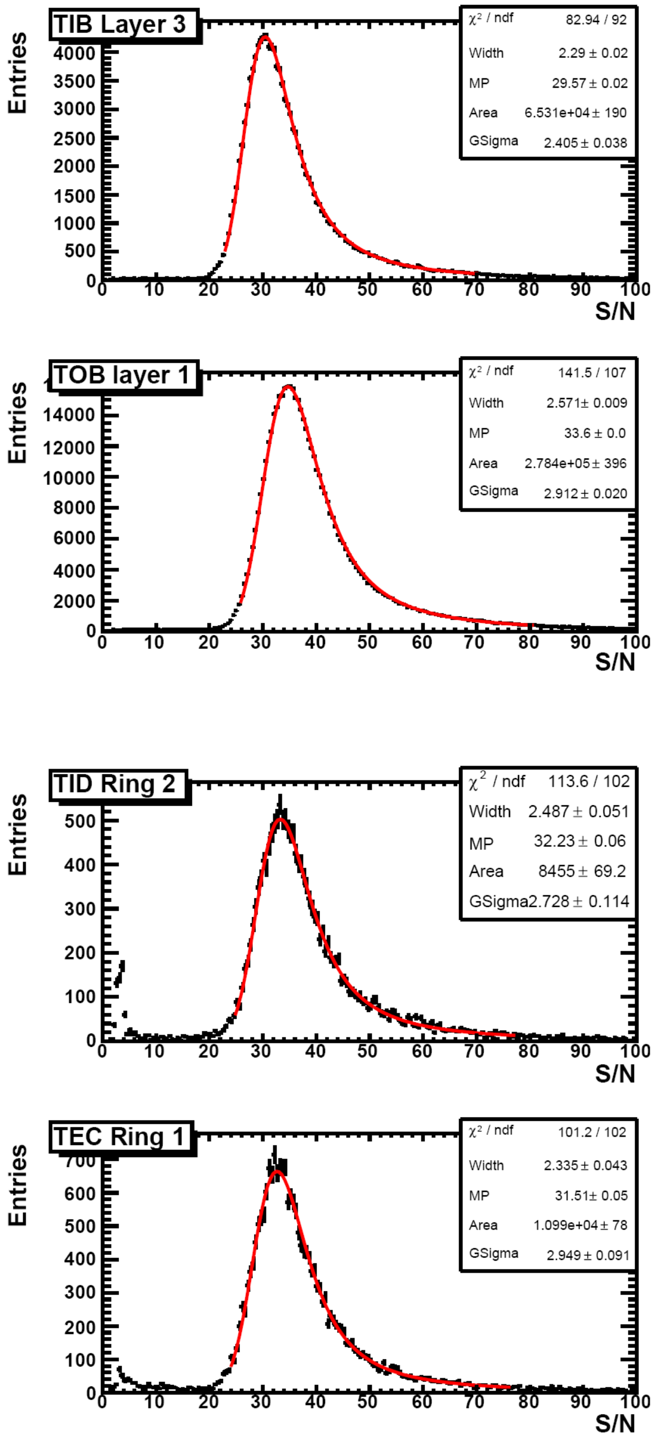


Figure 3.12: *The signal-to-noise corrected for the track angle for (from top to bottom) TIB L3, TOB L1, TID R2 and TEC R1 at T = -10°C.*

Layer 3, TOB Layer 1, TID Ring 2 and TEC Ring 1, while the full set of results from the fits of S_{ren}/N at T = -10°C are reported in Appendix in

Fig. 1 for TIB, TOB and in Fig. 2 for TID, TEC. The $\chi^2/d.o.f.$ values for the fits are excellent. For TEC and TID the statistics are lower than for TIB and TOB, since the rate of large angle cosmic tracks is much reduced with respect to the vertical ones. All tracker layers show a large S_{ren}/N, in all cases larger than 26.

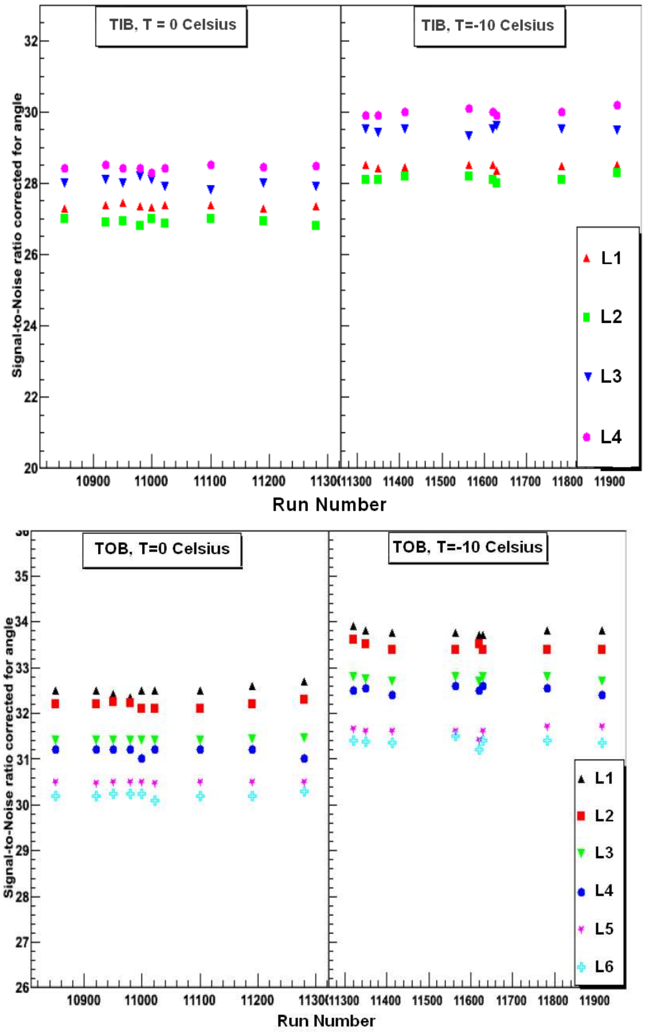


Figure 3.13: *Signal over noise corrected for the track angle: for TIB(top) and TOB (bottom) for two different temperatures.*

A study of the stability of S_{ren}/N performance was also done, examining measurements for every single run during long periods where the tracker was running in stable conditions: the best periods are of course the ones during which the tracker was running at low temperature. Fig. 3.13 shows the results for all the layers of TIB and TOB versus the run number corresponding to temperature of 0°C and -10°C. The two subdetectors show a very stable behavior with variations smaller than 0.3%. Similar results are obtained for TEC and TID, but the lower statistics per run gives rise to a higher statistical error on the run by run measurement.

S_{ren}/N increases by decreasing temperature, as expected from the results on the temperature dependence of the noise. A more quantitative analysis of the temperature dependence was not possible since it requires an omptimization of the APV parameters for each module at each temperature.

After the normalization of the signal to the sensor thickness, in principle the S_{ren}/N should be identical for identical modules, regardless to their position. Nevertheless, the segmentation of the silicon microstrips, combined with the effect of the clustering algorithm, which uses thresholds to determine strips inclusion in a cluster, might introduce some charge loss. The smaller the released charge, the more significant the fractional charge loss due to clustering threshold. This is particularly true for cosmic rays, which do not come from the interaction point and therefore produce tracks not according to the natural geometry of the tracker. A study of this effect has been done based on the angle of incidence of the tracks, as at higher angles charge tends to be shared among more strips. Therefore the dependence on the track angle has been studied first with respect to the 3D angle. In order to separate the effect of illuminating more strips from that of increasing the charge, the study has been done for: the dependence on the XZ local angle for the tracks with direction in a plane perpendicular to the silicon strips (YZ local angle less than 6 degrees), referred to as transverse tracks, where the sharing among strips is maximized; the dependence on the YZ local angle for tracks that come in a plane along the direction of the silicon strip (XZ local angle less than 6 degrees), referred to as longitudinal tracks, where sharing among strips is almost independent on the angle. These studies have been performed by looking at the dependence in S_{ren}/N on track path (K) within the silicon, where the path is expressed in units of silicon thickness. S_{ren}/N should be independent of K, as stated in equation 3.2.

Results obtained are shown in Fig. 3.14 for TIB and TOB, with black circles for the YZ angle and with open triangle for the XZ angle dependence. TEC and TID results have not been included since the variation of S_{ren}/N for the different rings is large and there is a correlation between K and the statistics for each ring.

It is evident in Fig. 3.14 that for small K there is a nearly 5% loss of S_{ren}/N (and presumably signal) for TIB. In the case of longitudinal tracks, most of the charge is released in one strip, and signal in neighboring strips increases with K: once above the clustering threshold, they are included in the cluster and therefore almost all the charge released by the track is taken into account already for $K = 1.3$, where a sudden increase is visible. On the other hand, in the case of transverse tracks, the charge is shared among several strips and the charge loss at the edges of the cluster will be less significant as K increases. The transverse and

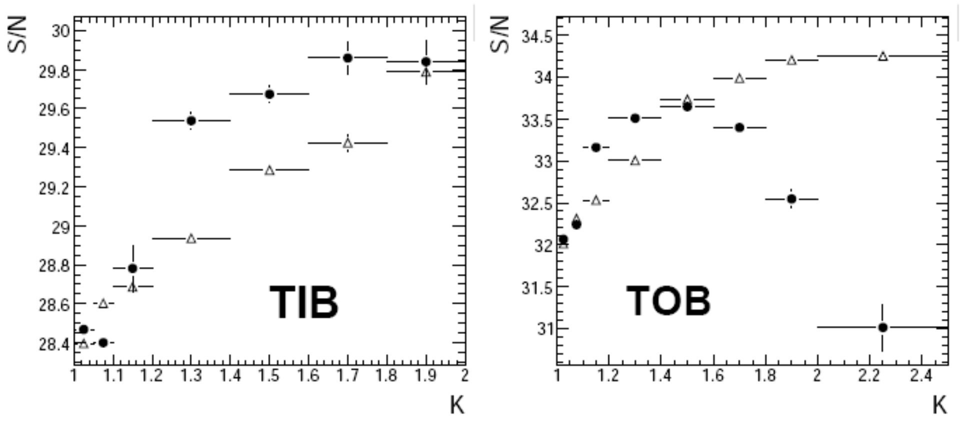


Figure 3.14: *S_{ren}/N dependence on K for tracks perpendicular to the strip direction (open triangles), i.e. the XZ local direction, and for tracks parallel to the strip direction (black circles), i.e. the YZ local direction, for TIB and TOB*

parallel tracks give the same signal-to-noise value for $K = 2$.

TOB shows a similar behavior, with a nearly 6% estimated loss of charge for small K. The charge released by particles with transverse tracks increases almost linearly with K and parallel tracks show a sudden increase at $K = 1.2$. Contrarily to the TIB, TOB does not reach a constant value for parallel tracks, but S_{ren}/N decreases starting from $K = 1.6$. This effect can be interpreted considering the effect of departure from linearity of the APVs for charges higher than 3 mips (minimum ionizing particles) with gradual fall off beyond, that influences the tail of the Landau distribution for large charge release to single strips.

3.3.2 The offline gain calibration

The correction for the electronic gain as shown in Sec. 2.3.1 can be used to improve the resolution on the signal performance. Fig. 3.15 shows the signal before and after the correction. For all sub-detectors the calibration results in a decrease of the FWHM; it is clear that the electronic gain calibration improves the FWHM. In TEC, after the electronic calibration, it is even possible to distinguish the two peaks, as a consequence of the use of thin and thick sensors.

Including this calibration permitted to check the level to which the absolute calibration is understood.

To do this, an analysis at the module level has been done, looking at the distribution of the most probable values of the signal, expressed in ADC counts, for the thin detectors (TIB and TEC rings from 1 to 4) and for the thick detectors (TOB and TEC rings from 5 to 7) separately. Results are shown in Fig. 3.16 for the -10°C period.

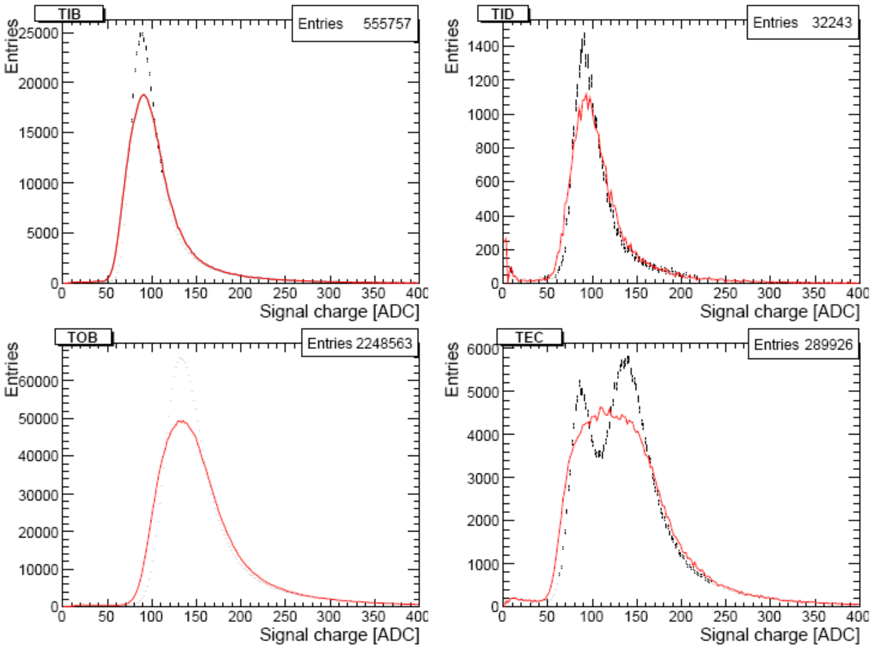


Figure 3.15: *Signal corrected for the track angle (S_{ren}), in full curve without electronic gain calibration and in dots with it.*

Modules of the same subdetectors distribute in Gaussian shapes with fairly large sigmas at a level of 3.8% and 5.3% for TIB and TEC thin sensor modules, and of 4.6% for TOB and TEC thick sensor modules. Therefore, the signal performance of single modules can be understood to these levels of precision within the same sub-detector, to be compared with the spread of the tickmark height of about 13% before taking into account the electronic gain. Further improvements in the evaluation of the electronic gain can come, for example, by taking into account the supply voltage for individual modules.

It should be noted also that modules of the same geometry but placed in different positions are differently illuminated by the cosmic rays and this causes variations in S_{ren}, as it has been illustrated in last section. This constitutes another level of complexity in the study.

Some lessons can also be learned by comparing the average results obtained by silicon modules of the same thickness: for the TIB and thin sensors TEC modules, the average values are within 2.2%, while TOB and thick sensors TEC modules have average values that differ by 5%. It should be borne in mind that signal performance can be affected by changes in APV parameters. Typically, these are adjusted for reasons other than equalizing the peak amplitude of the output signal.

The TEC has the interesting feature that all modules, both thin and thick, have been run with the same APV parameters. The ratio of the

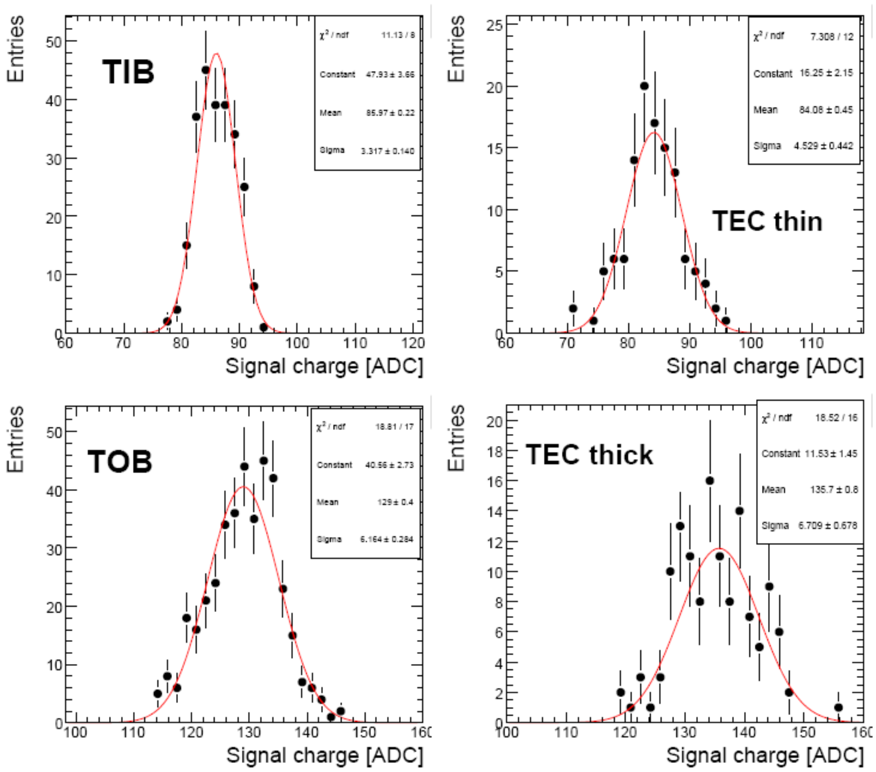


Figure 3.16: *Distribution of signal (S_{ren}) most probable value, module per module for TIB (top left), TEC thin sensors (top right), TOB (bottom left) and TEC thick (bottom right).*

mean values of thick over thin sensor module signal is: $\frac{<Signal(TEC_{thick})>}{<Signal(TEC_{thin})>}$ = 1.614±0.008, which is compatible with the ratio 1.62 of thicknesses of 470 μm and 290 μm. This is not so precisely the case when TIB and TOB are compared.

In order to better understand the signal performance, this analysis has been performed at two different temperatures: 10°C and -10°C. The signal changes of 2-3% within a given layer for the 20 degrees change in temperature. It can be concluded that the conversion factor between ADC counts and electrons must be done at individual layer level for each temperature if a 5% level has to be reached. Further refinements to the calibration would require a module by module study based on track information. This has been done later, during the Global Runs (Sec. 3.8.1).

3.3.3 Cluster width studies

In silicon sensors the electric field is normal to the strips and therefore, in absence of magnetic field, the drift direction of charge carriers is coincident with the field lines. Hence for normal incidence particles only one strip is hit, and the cluster size increases with the incident angle. In the presence of a magnetic field, the drift direction is no longer along the electric field drift lines, as shown in Fig. 3.17. The diffusion path of charge carriers inside the bulk silicon is influenced by the Lorentz force, which hence modifies the shape of the charge distribution of the clusters. The net effect is a shift in the cluster position (relative to track intersection point) and a change of the cluster width. The shift and widening are largely due to charge drifting across cell boundaries, but the capacitive coupling between neighboring channels also has an influence on the final position and cluster width.

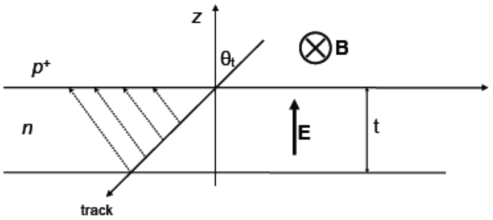


Figure 3.17: *Lorentz drift in the micro-strip sensors.*

In the case of the Sector Test, where there was no magnetic field, it is still interesting to evaluate the behavior of the cluster width as a function of the track incident angle perpendicular to the strip direction (angle XZ described in Fig. 3.11).

Since the value of the Lorentz angle might change as the silicon is irradiated a calibration strategy has been developed which will determine and monitor its value from data and write it in a database, so that it can be retrieved by the offline analysis code, to correct the clusters position. TIF datasets have been used to validate and test this software with real data for the first time.

The average cluster width for tracks incident with an angle θ_t with respect to the detector normal is given by:

$$< clusterwidth >= a + |\frac{t}{p} \cdot b \cdot (\tan\theta_t - \tan\Theta_L)| \tag{3.3}$$

where t is the detector thickness, p is the pitch, a and b are coefficients expressing the carrier diffusion and the electronic capacitive coupling between nearby channels, and Θ_L is the Lorentz angle. Therefore, plots of mean cluster width as a function of tan θ_t are produced for each layer and the Lorentz angle is obtained fitting them with the function above. In this fit the parameters are: the estimate of the Lorentz angle (tan Θ_L), the slope normalized to the ratio of thickness over pitch (that should be constant), and the average cluster width at the minimum. Of course in the TIF setup there is no magnetic field and the Lorentz angle determination is expected to return zero.

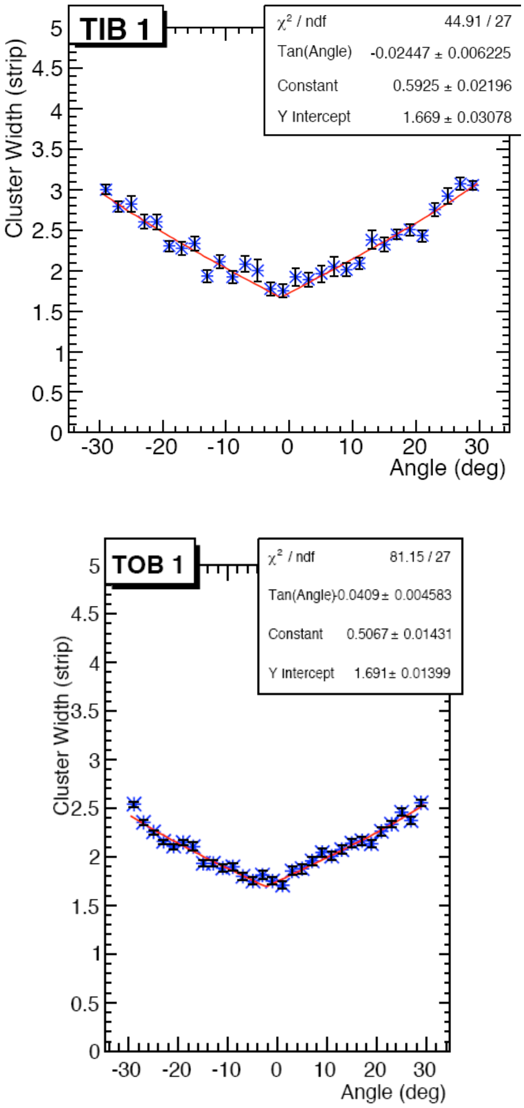


Figure 3.18: *Cluster width versus* θ_t *for TIB layer 1 (top) and for TOB layer 1 (bottom).*

The results are shown in Figs. 3.18 for TIB Layer 1 and TOB Layer 1 and in Appendix Fig. 3 and 4 for all layers. The fitted Lorentz angles are consistent with zero as expected within the statistical uncertainties, and with a disagreement of 2% at the most in some cases likely due to systematic misalignment effects.

3.3.4 Hit occupancy

Occupancy for a given strip is defined as the fraction of events in which it registers a signal exceeding the threshold required to be a “hit”. When LHC will run at high luminosity, the occupancy of the innermost strip layer (TIB layer 1) is estimated to be a few percent [54]. As described in Sec. 2.4.1, seed strips and the cluster require a signal-to-noise ratio value greater than 3 and 5 respectively. If strip noise is purely Gaussian then the probability of a statistical fluctuation to become a cluster is very low, less than 5×10^{-5}, and is therefore negligible. The subdetector occupancy has been measured during Sector Test by taking the mean value of the distribution of the probability of a strip to be in a cluster, that is, the strip occupancy distribution.

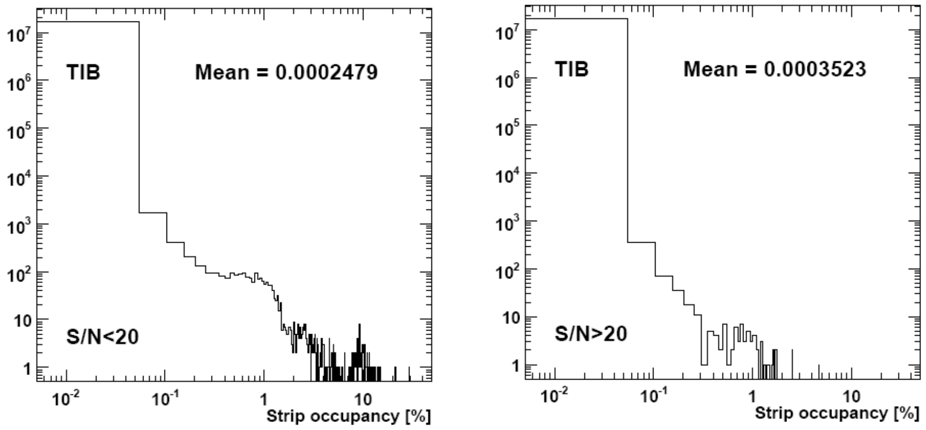


Figure 3.19: *Distribution of strip occupancy for TIB: left (right) plot for S/N below (above) 20.*

In order to separate the contribution from real tracks, the strip occupancy has been studied for clusters with S/N higher or lower than 20. Results are shown in Fig. 3.19 for TIB. Cosmic rays account for an occupancy of 3.5×10^{-6} with no tails in the distribution. Clusters from statistical fluctuation accounts instead for an occupancy of 2.5×10^{-6} and show much longer tails. It is evident that there are strips more active than others, with occupancies up to 5% or larger.

It seems clear that the tails observed in some of the occupancy distributions are the result of active (“hot”) strips. These strips have not to be confused with the noisy strips described in Sec. 3.2.3, since these “hot” strips are characterized by non-Gaussian fluctuation, above five times the noise value. The number of “hot” strip occupancy above 1% is below 0.05% and therefore does not significantly affect module occupancy. Similar results have been obtained for TOB, TEC and TID subdetectors.

3.3.5 Hit efficiency

The hit efficiency is the probability to find a cluster in a given silicon sensor that has been traversed by a charged particle. In order to calculate the hit efficiency, track seeding, finding and reconstruction must be performed. The results presented here have been determined using the Combinatorial Track Finder for cosmic muon events introduced in Sec. 3.1.3.2, excluding the clusters in the layer of the tracker for which the hit efficiency is to be determined. The efficiency for a given module in this layer is then calculated by finding tracks that pass through that module and determining if a cluster is, in fact, present.

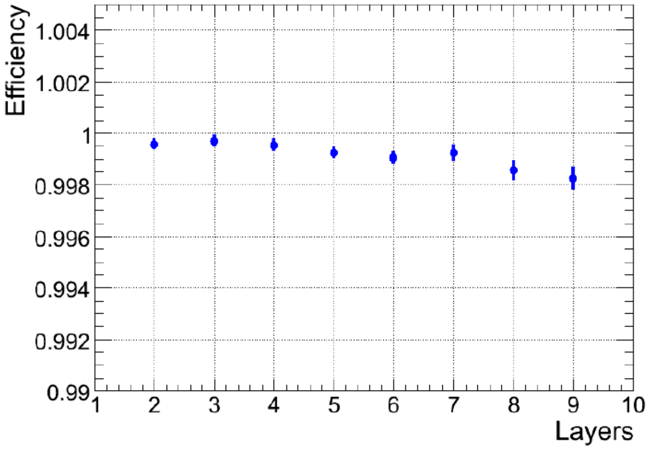


Figure 3.20: *Layer efficiency at T = -10° C. Layer 1 (TIB layer 1) and layer 10 (TOB layer 6) are missing because of the definition of the method: to measure their efficiency, an external layer would have been necessary to trigger.*

A sample of high quality events was selected by requiring only one track reconstructed, one hit in the first TIB layer, one hit in the two outermost TOB layers, and at least four reconstructed hits (of which at least three matched from stereo layers). Tracks were required to have no more than five lost hits during the pattern recognition phase and not more than three consecutive ones; finally, an upper cut of 30° on θ_{3D} was applied to select tracks almost perpendicular to the modules.

In order to avoid genuine inefficiencies present at the edge of the sensor and in the bonded region between two sensors, additional cuts have been applied to restrict the region in which efficiency is measured. With this event selection criteria and fiducial area restrictions, the efficiency exceeded 99.8% for all measured layers as is shown in Fig. 3.20.

3.3.6 Tracking performance

To visualize the standard track reconstruction results, the data sample taken at -10°C in trigger configuration C using all silicon strip subdetectors has been chosen.

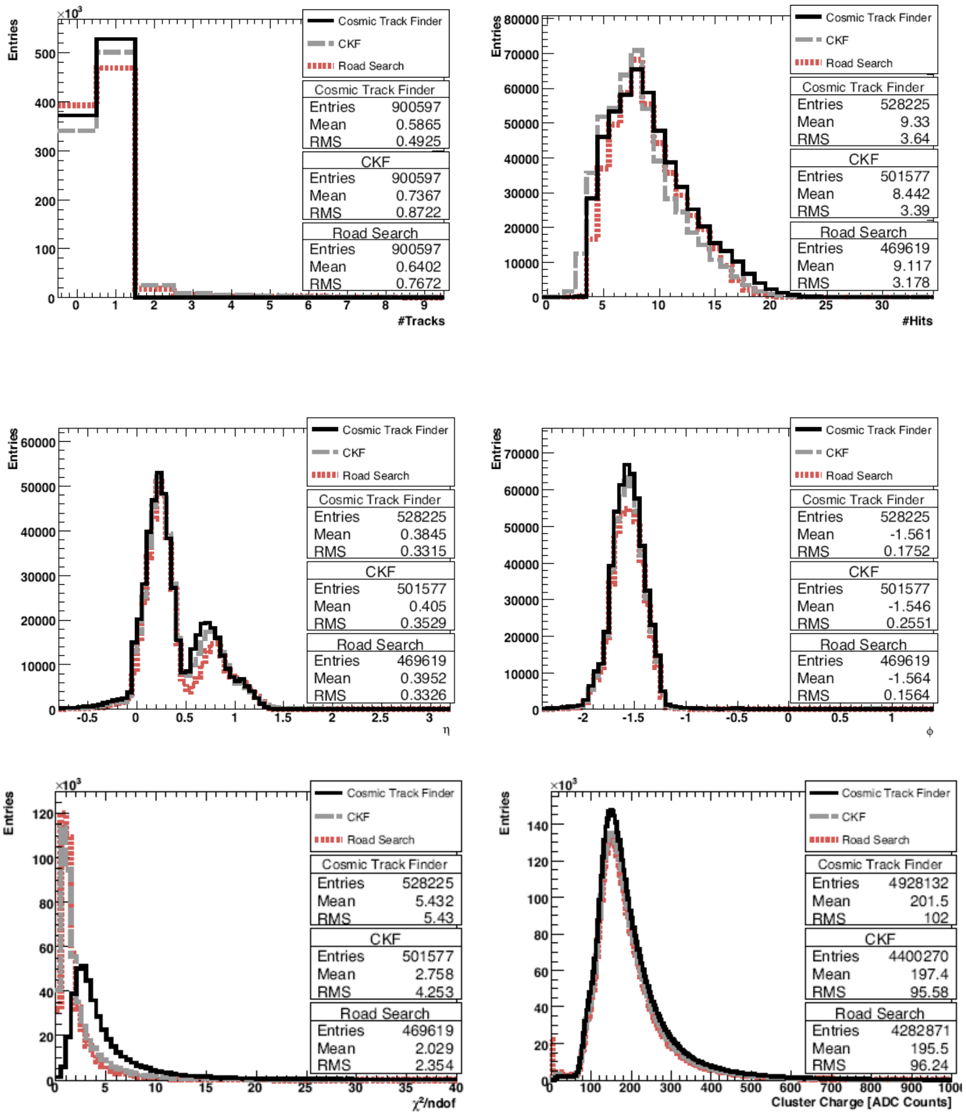


Figure 3.21: *Number of reconstructed tracks and various track distributions in single track events using data taken at T = -10°C in trigger configuration C with TIB+TOB+TEC as active detector.*

The number of reconstructed tracks for all three tracking algorithms is shown in Fig. 3.21, along with various track distributions in single track events. Apart from the different numbers of reconstructed tracks, expected given the capability of the Cosmic Track Finder to reconstruct only one track per event, all three tracking algorithms lead to similar results. The ϕ distribution shows a peak around -$\pi/2$, compatible with tracks that originate from the top of the detector and travel outside in. The η distribution is compatible with the trigger layout. The number of $r\phi$ and stereo hits per track shows a most probable value of 8 hits for all algorithms.

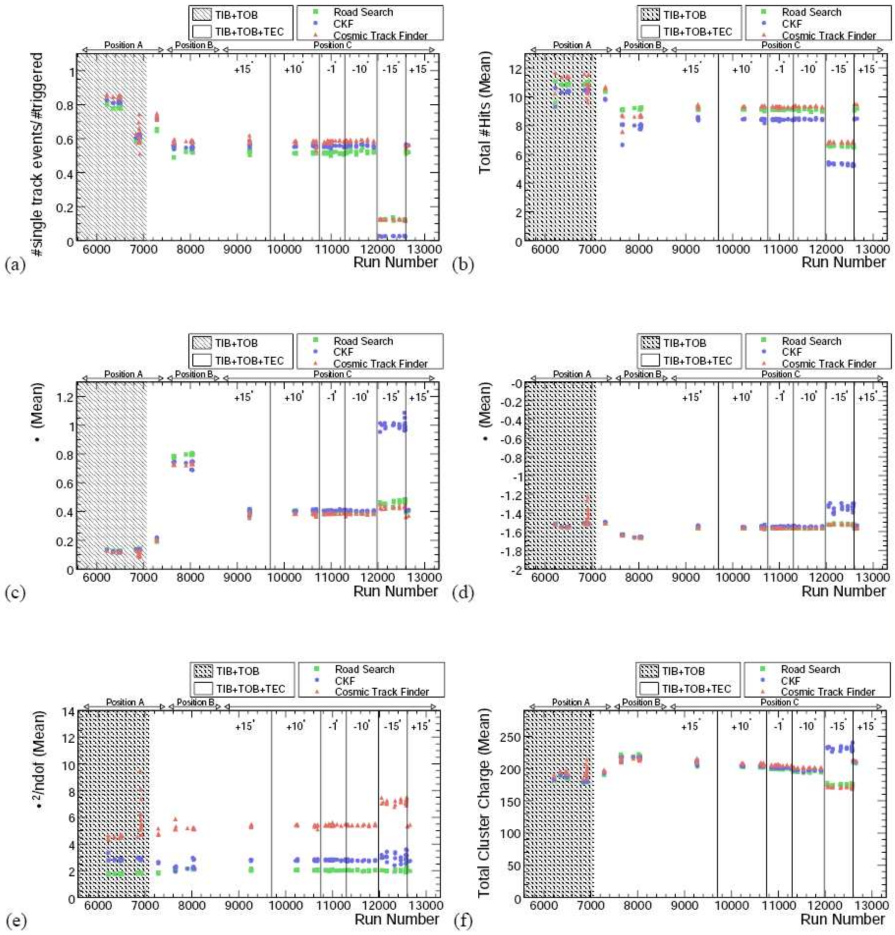


Figure 3.22: *Control plots showing the mean of various reconstructed track parameters for all three algorithms as a function of run number. The scintillator positions and the operating temperatures (in °C) are indicated at the top of each frame. All distributions are shown for single track events: (a) number of single track events divided by triggered events; (b) number of hits; (e) χ^2/ndof; (f) cluster charge in ADC counts of hits belonging to the track before gain correction. The differences observed at -15°C are due to the reduced number of active modules.*

Control plots have been generated in order to monitor the stability of the track reconstruction results and to study the dependence on the operating temperature as a function of time. Figure 3.22 shows the distribution of various reconstructed track parameters for single track events from all three algorithms. For each cosmic run, the mean of the parameter under study has been extracted and is plotted as a function of the run number. Each histogram indicates the active detectors, the various scintillator positions and the respective operating temperature. It is expected that the tracking parameters do not depend on the operating temperature, while a clear dependence on the trigger setup and the active detectors should be seen. This can be observed for trigger configuration C, where the temperature was gradually decreased from +15°C to -15°C. The reconstructed track parameters do not show a dependence on the operating temperature, except for data taken at -15°C. In order to reach this operating temperature, a significant amount of modules in all four sub-detectors had to be turned off, effectively changing the active detectors. The tracking algorithms were not retuned for this situation. The differences in their seeding configurations translate into a different acceptance and explain the variations in the number of reconstructed tracks and the track parameter distributions.

3.3.7 Track reconstruction efficiency

TIF data were used to measure for the first time the efficiency of the tracking algorithms on a fraction of the tracker. The estimation of the tracking efficiency is a challenging task, due to the poor prior constraints available for cosmic tracks and the absence of an external reference other than the scintillator counters. The strategy adopted for this analysis is the reconstruction of partial tracks, using only a subset of the tracker. These track segments serve as a reference for the reconstruction in the remaining parts of the tracker. For statistical reasons, tracks in the TIB and TOB were chosen as independent subsets of the tracker. In the following, track segments reconstructed in TIB will be referred to as TIB tracks, while track segments reconstructed in the TOB are referred to as TOB tracks.

To isolate single track events, only events with less than 30 reconstructed hits were analyzed. TIB and TOB tracks were accepted if the normalized χ^2 was smaller than 30 and if they contained hits with signal-to-noise ratio greater than 8 in at least four different layers. Based on the selected events, the tracking efficiency in TIB (TOB) was calculated from the fraction of TOB (TIB) tracks with a matching track in the other sub-detector. This conditional efficiency can deviate from the global efficiency if the TIB and TOB acceptance differ after selection. The match between tracks was based on a comparison of the azimuthal angles. The difference was required to be smaller than five times the resolution determined from simulation. In addition, the reference track had to contain at least two hits in stereo layers and the extrapolation had to be fully contained in the active region of the other sub-detector.

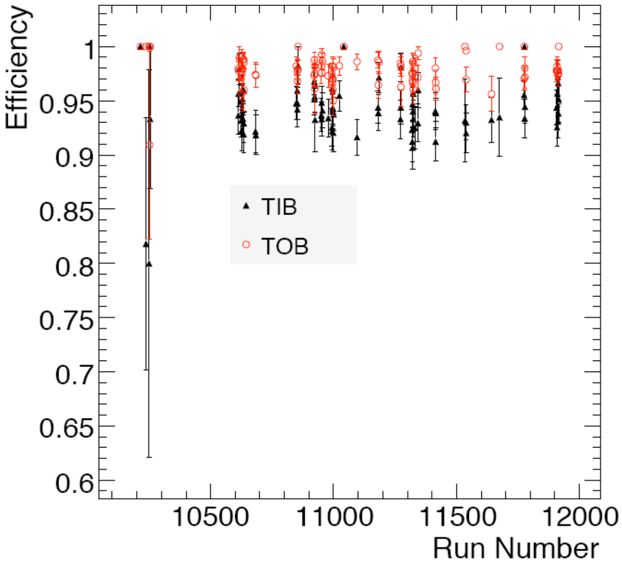


Figure 3.23: *Conditional track reconstruction efficiencies* ϵ*(TIB|TOB) (black triangles) and* ϵ*(TOB|TIB) (red circles) for the CKF as a function of the run number.*

	ϵ(TIB|TOB) [%]		ϵ(TOB|TIB) [%]	
	Data	MC	Data	MC
CKF	94.0±0.2	98.66±0.04	97.7±0.1	98.76±0.04
CosmicTk	93.1±0.2	94.46±0.09	96.9±0.1	97.36±0.06
RS	89.9±0.2	89.08±0.12	99.0±0.1	99.39±0.03

Table 3.5: *Average conditional track efficiencies and corresponding statistical uncertainties for the three track reconstruction algorithms in data and Monte Carlo simulation.*

The average efficiencies for all three track reconstruction algorithms for data and Monte Carlo simulation are shown in Table 3.5. The results have been obtained using data taken at +10°C, -1°C and -10°C in scintillator position C. The stability of the conditional track reconstruction efficiency is demonstrated in Figure 3.23, which shows the distribution of ϵ(TIB|TOB) and ϵ(TOB|TIB) for the CKF as a function of run number.

3.4 The Tracker alignment test

Accurate studies have been performed at TIF with the algorithms descried in Sec. 2.6.3 in order to determine the maximal set of detectors that can be aligned and the aligned coordinates that are sensitive to the peculiar track pattern and limited statistics of TIF cosmic track events.

In addition also a third statistical method have been applied, the Kalman alignment algorithm [74]. This is a sequential method, derived using the Kalman filter formalism. It is sequential in the sense that the alignment parameters are updated after each processed track. The algorithm is based on the track model $m = f(q_t, p_t) + \epsilon$. This model relates the observations m to the true track parameters q_t and the true alignment constants p_t via the deterministic function f. Energy loss is considered to be deterministic and is dealt with in the track model. The stochastic vector ϵ as well as its variance-covariance matrix V contain the effects of the observation error and of multiple scattering. Therefore the matrix V contains correlations between hits such that equation 2.3 is a sum over tracks, with residuals being of higher dimension according to the number of hits along the track trajectory. Linearised around an expansion point (q_0, p_0), i.e. track parameters from a preliminary track fit and an initial guess for the alignment constants, the track model reads:

$$m = c + D_q q_t + D_p p_t + \epsilon \tag{3.4}$$

with

$$D_q = \partial f / \partial q_t|_{q_0}, D_p = \partial f / \partial p_t|_{p_0}, c = f(q_0, p_0) - D_q q_0 - D_p p_0 \tag{3.5}$$

By applying the Kalman filter formalism to this relation, updated equations for the alignment parameters p and their variance-covariance matrix C_p can be extracted.

For the tracker barrels (TIB and TOB), the collected statistics has been sufficient to align at the level of single modules if restricting to a geometrical subset corresponding to the positions of the scintillators used for triggering. The detectors aligned are those whose centres lie inside the geometrical ranges $z > 0$, $x < 75$ cm and $0.5 < \phi < 1.7$ rad where all the coordinates are in the global CMS frame.

The local coordinates aligned for each module are

- u, v, γ for TOB double-sided modules,
- u, γ for TOB single-sided modules,
- u, v, w, γ for TIB double-sided modules and
- u, w, γ for TIB single-sided modules.

Due to the rapidly decreasing cosmic track rate proportional to $\cos^2\Psi$ (with Ψ measured from zenith) only a small fraction of tracks crossed the endcap detector modules at an angle suitable for alignment. Therefore, the z^+-side Tracker endcap (TEC) could only be aligned at the level of disks. All nine disks were considered in TEC alignment, and the only aligned coordinate was the angle $\Delta\phi$ around the CMS z-axis. Because there were data only in two sectors of the TEC, the track-based alignment is not sensitive to the x and y coordinates of the disks. The Tracker Inner Disks (TID) were not aligned due to lack of statistics. Fig. 3.24 shows

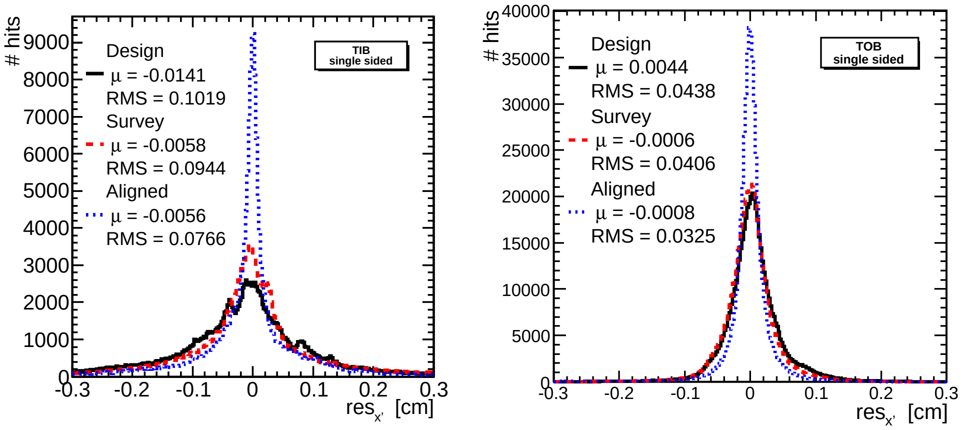


Figure 3.24: *Hit residuals for different geometries: ideal (solid/black), survey (dashed/red), and track-based alignment (dotted/blue, HIP). Two Tracker sub-detectors are shown: TIB in the left figure, TOB in the right one. The residuals for the remaining subdetectors are shown in Appendix, Fig. 5.*

the hit residuals in TIB and TOB in the design geometry, after the survey measurements, and after the alignment with the track-based algorithms. Both the overall track quality and individual hit residuals improve in the three steps. Fig. 3.25 shows the evolution of the χ^2 from the design geometry to the aligned one.

All the track-based algorithms produce similar results when the same input and similar approaches are taken, as shown in Fig. 3.26 for TOB as an example, and in Appendix in Fig. 6 for all the other subdetectors. The distributions show the hit residuals with respect to the track, for different geometries from the three track-based algorithms. The comparison between the results is an important validation of the systematic consistency of the methods. It can also be shown that the residual misalignments are consistent with statistical uncertainties in the procedure.

The typical achieved precision on module position measurement in the local x coordinate is estimated to be about 50 μm and 80 μm RMS in the Tracker Outer and Inner Barrels, respectively. However, since no magnetic field was applied in the tracker, no momentum estimate of the cosmic tracks was possible. Therefore, detailed understanding of alignment precision suffered from uncertainties in multiple scattering of tracks with unknown momentum, this being the dominant contribution to

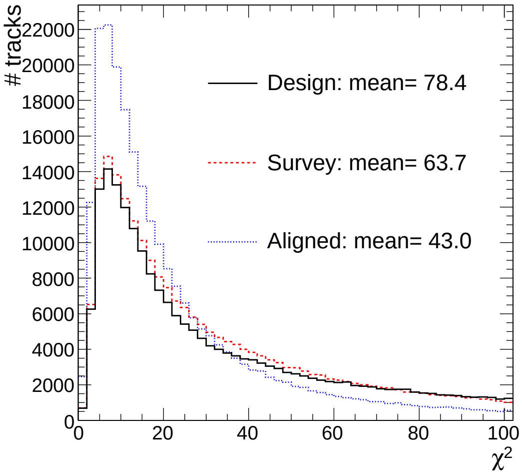


Figure 3.25: *Distribution of the absolute χ^2-values of the track fits for the design and survey geometries as well as the one from HIP track-based alignment.*

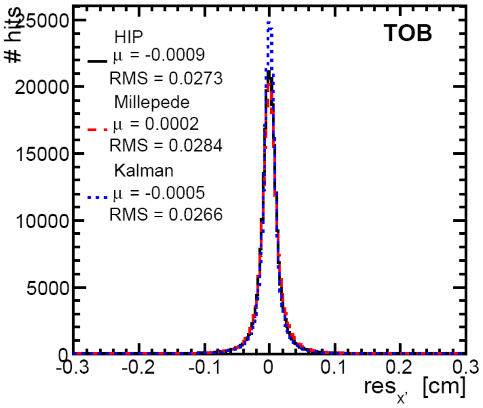


Figure 3.26: *Hit residuals in TOB from three track-based algorithms: HIP (solid/black), Millepede (dashed/red), and Kalman (dotted/blue) based alignment. The absolute local x'-residuals are shown for single-sided modules.*

the hit position extrapolation. A more accurate test has become possible when the Tracker has been inserted in CMS and has taken data with magnetic field switched on, during CMS Global Runs.

3.5 Summary on Sector Test

The Sector Test has been an excellent moment to test not only the detector itself, but also the reconstruction and calibration algorithms, the Data Quality Monitoring and the analysis workflow. The 4.5 million cosmic ray data collected have been used in all the possible ways to permit a full comprehension of the Tracker, even if only a portion of it was operative. In this portion, the percentage of dead channel has been found to be very low, much less than 1%. The Tracker behaved very stably, with a signal-to-noise ratio larger than 26 for all layers and rings.

The Lorentz angle has been found to be zero in absence of magnetic field, as expected, and a first gain calibration has been done.

The hit occupancy has resulted to be of the order of 10^-6 and the hit efficiency showed a very high value of 99.8%.

The tracking efficiency in the small available fraction of the detector has been measured to be larger than 93% for $\epsilon(\mathrm{TIB}|\mathrm{TOB})$ and larger than 96.9% for $\epsilon(\mathrm{TOB}|\mathrm{TIB})$ for both the Combinatorial Track Finder and the Cosmic Track Finder.

A preliminary attempt to align the Tracker has been also done, and the algorithms have been tested, giving a final alignment of about 50-80 μm as mean of the residuals, which is a very good result considering the absence of the magnetic field and therefore the impossibility to measure the momentum.

All this experience has been then used later, during the test of the full Tracker within the whole CMS detector, described in the next section.

3.6 The Tracker in the CMS Global Runs

After the test performed at TIF on a small fraction of the whole tracker, the silicon detector has been moved from the Integration Facility to its final location at the CMS site, and integrated into the experiment with all the other subdetectors starting from December 2007. It was assembled and installed in the CMS detector until spring 2008 and finally commissioned in summer 2008.

During Global Runs of CMS, cosmic rays were detected in the muon chambers and used to trigger the readout of all CMS sub-detectors. In this first phase the magnet was still off, and the test was called Cosmic Run at Zero Tesla (CRuZeT). Afterwards, the superconducting solenoid provided a 3.8 T magnetic field. This test has been called Cosmic Run at Four Tesla (CRAFT). I will focus in particular on this second phase, describing all the studies done and the results obtained.

3.6.1 Data Sets

The whole CRAFT data-taking period can be splitted in several subsets as shown in Table 3.6 according to tracker and magnetic field conditions. In the first period part of the tracker was not correctly synchronized with the rest of CMS. The reason was eventually traced to a wrong fiber lenght value used during the synchronization and a new latency scan fixed the problem. Then, after a period of data-taking with stable magnetic field at 3.8 T, some data were taken with magnet off.

Period	**First Run**	**Last Run**	**Events/10^6**
A	66604	69256	10.9
B	69269	70088	2.2
C	70195	70675	1.6

Table 3.6: *Subsets of data with uniform conditions. Periods B and C have similar signal and noise conditions, but in period C the magnetic field is off. The number of events refers to tracker pointing selection.*

For tracker studies a selection of data is applied, here called tracker pointing: the selected events must have at least one reconstructed track in the tracker, or have a particle reconstructed in the muon chambers pointing to a cylinder with 90 cm radius and 260 cm length modeling the tracker. In total more than 2 million events were recorded in the tracker pointing selection in period B. The data quality was monitored during online and offline reconstruction and propagated to the CMS DataSet Bookkeeping System which indexes event-data for the CMS Collaboration. Unless stated otherwise only good runs taken at B = 3.8 T stable field are used for the analyses presented in my thesis.

In addition to data, several analyses use a simulated sample of cosmic muons to compare results. The sample was generated in similar way to the sample generated for comparison with TIF data. The difference is in the filter before the detector simulation, which in this case selects only those events with a cosmic muon in the fiducial tracker volume. Furthermore, only muons with momentum greater than 4 GeV/c were kept, as this corresponds to the minimum momentum required to penetrate through the tracker. The final simulation sample contains more than 21 million events.

3.6.2 Data Quality Monitoring

During the period between the TIF data taking the DQM for the Tracker has been intensively developed to meet the analysis requirements. I was the responsible for one of the developments to increase the powerfulness of the system for the monitoring of the data quality. I developed a DQM module for the clusters properties monitoring, both when they are associated to a track and when they are noise clusters. This module is the natural evolution of the analysis program I illustrated in the previous section, but with the fundamental difference that, being integrated in the DQM, it makes possible to monitor the quantity related to the clusters directly online, and this of course permit to spot possible problems very soon and to react promptly.

Moreover, thanks to the offline DQM, it is still possible to perform the same analysis also offline, perform fits and study the trends in long periods. The results that will be shown in Sec. 3.8.2 has been obtained using the output of this DQM module.

3.7 Commissioning and calibration

The different phases of the tracker commissioning have been introduced in Sec. 2.3. Here the results of these calibration procedures before starting the CRAFT data taking are shown.

3.7.1 Preliminary gain calibration

The first calibration regards the gain and the AOH optimization. Fig. 3.27 shows the tick-mark height distribution after the AOH optimization. The values are close to the planned 640 ADC counts and the spread is due to the rough setting capabilities for the gain (1, 1.5, 2, 2.5).

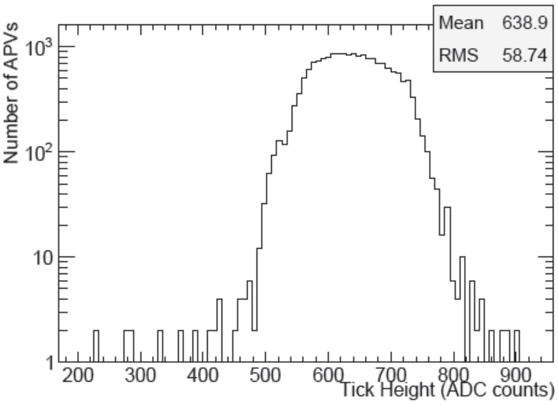


Figure 3.27: *TOB tick-mark height after AOH optimization.*

The residual miscalibration has been corrected offline with the calibration of the signal. The result will be reported in Sec. 3.8.1.

3.7.2 Synchronization to an external trigger

The main difference of the Global Run with respect to the data taking at TIF is for sure the need to synchronize each tracker partition with an external trigger, which is the Level 1 trigger of CMS. The trigger needed for local runs was provided by the Drift Tube technical trigger, a coincidence of central top and bottom chambers.

Two procedures were applied to perform the synchronization: the latency scan, and a fine delay scan.

3.7.2.1 Latency scan

The first procedure is a coarse scan in time, by steps of 25 ns, by adjusting the latency between the trigger arrival and the sampling time of the APV chip. The mean signal of the leading strip (the channel with the largest signal amplitude) in clusters associated to reconstructed tracks was extracted as a function of the latency. The signal magnitude was corrected for the track path length through the active sensor volume, inferred from the track angle.

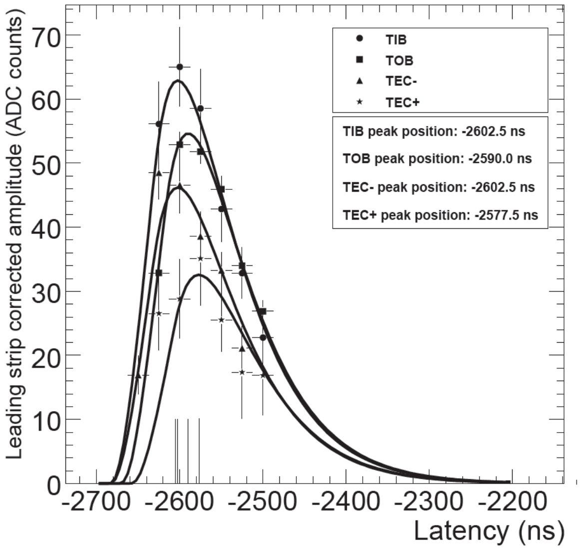


Figure 3.28: *Mean signal of leading strip in clusters associated to tracks as a function of the latency, for each of the four partitions.*

The measurement was performed for the tracker as a whole (rather than individual partitions). Unfortunately, it was discovered post data-taking that an incorrect trigger cable length for TOB was used in the calculation. This resulted in adjustments to the hardware configuration such that the front-end modules in the TOB received an out-of-phase clock signal, shifted by 12.5 ns with respect to the three other partitions. Since the statistics collected in the TOB dominated the measurement, the adjustments were done relative to the TOB results. TIB and TEC- were shifted by 12.5 ns and TEC+ by -12.5 ns, as shown by the fits in Fig. 3.28. Time-of-flight is not taken into account in this procedure, since the variations expected across the detector ($\leq$10 ns with cosmics, 5 ns in collisions) are lower than the range of precision of 25 ns.

3.7.2.2 Fine delay scan

The last procedure comprises a fine tuning of the synchronization and was not performed until the very end of the CRAFT data-taking pe-

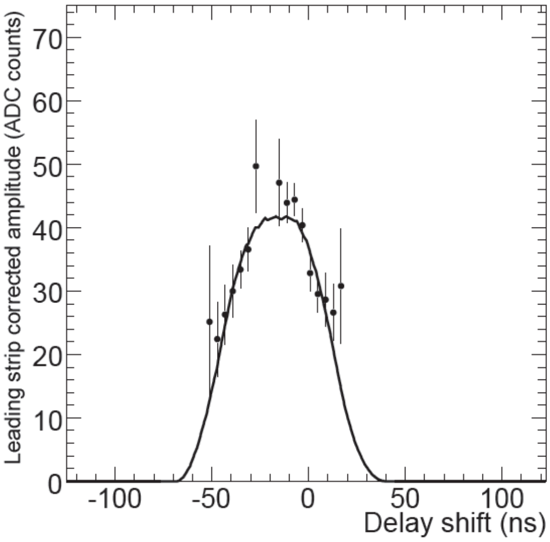


Figure 3.29: *Fine delay scan for the TOB layer 3, in deconvolution. The mean position (-14.2 ns) has to be compared to the mean time-of-flight of particles from the muon system to the silicon sensors (12 ns).*

riod. It consists in skewing the trigger and clock delay by steps of 1 ns around the expected optimal value for all modules of a given test layer, the other modules staying at the value obtained by the latency scan. Clusters on the test layer compatible with a reconstructed track are used to reconstruct the pulse shape.

Fig. 3.29 shows the resulting pulse shape from clusters found in module of TOB layer 3, acquired in deconvolution mode. With collision data, the time-of-flight can be adjusted for each individual track, but this is not the case for cosmics, for which the jitter from the trigger cannot be subtracted. Additionally, the origin of the tracks is less well defined, as there is no interaction point. The 14 ns shift observed is consistent with the expected time-of-flight (12 ns) of cosmic muons from the Muon Drift Tube chambers to the TOB layer 3.

From the analysis of latency and fine delay scans, correction factors can be computed to compensate the residual mis-synchronization of each partition. These factors are presented in Table 3.7. They correspond to the ratio of the amplitudes at the supposed working point and at the maximum of the CR-RC curve, and have to be used to correct the cluster charge in calibration studies.

Partition	Correction Factor
TIB/TID	$1.018\,^{+0.012}_{-0.009}$
TOB	$1.0013\,^{+0.0065}_{-0.0012}$
TEC+	$1.058\,^{+0.032}_{-0.023}$
TEC-	$1.018\,^{+0.012}_{-0.009}$

Table 3.7: *Correction factors from synchronization for each partition. The uncertainty accounts for a 3 ns resolution on the ideal sampling time and includes residual time-of-flight effects in the tracker volume.*

3.7.3 Noise studies and detector quality tests

The noise is calibrated with the measurement described in Sec. 2.3.2. The different wafer geometry are studied separately to take into account the various strip lengths and pitch adapters that are modifying the strip input capacitance. The noise of TOB layer 3 $r\phi$ modules, expressed in electrons, is shown in Fig. 3.30. Outliers correspond to missing high-voltage on some modules, leading to a much higher input capacitance.

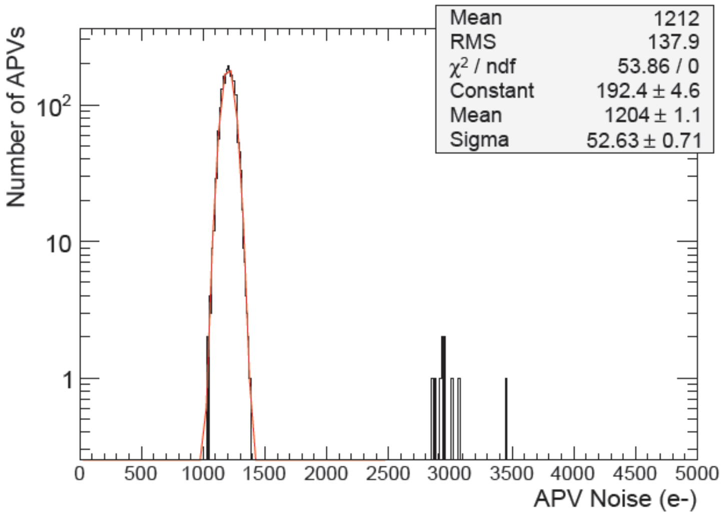


Figure 3.30: *Normalized noise measurement for all APVs of TOB $r\phi$ Layer 3 modules. Modules with HV OFF clearly appear.*

Fitting the noise versus the strip length, also in this case as in TIF data taking, a linear dependency is obtained:

$$N(e^{-}) = (427 \pm 39) + (38.7 \pm 3) \times L(cm) \tag{3.6}$$

As already mentioned, Kolmogorov tests can be performed on the noise distribution tails in order to spot non-gaussian (bad behaving) modules.

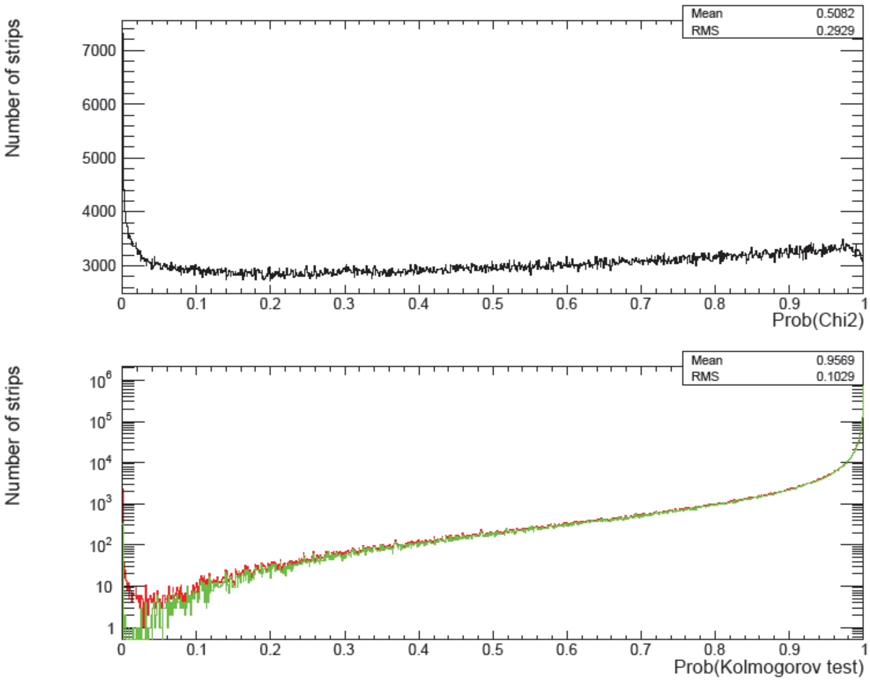


Figure 3.31: *The χ^2 probability of all strips of the TOB is shown. It cannot be used to build an efficient rejection cut. The bottom plots shows the Kolmogorov test results for all strips (red) and strips with Prob(χ^2) $\geq$ 0.1 (green). A 2% cut rejects all bad strips.*

The χ^2 and Kolmogorov probabilities for a fit of the noise with a Gaussian distribution obtained during Global Runs are plotted in Fig. 3.31. Very few strips (~0.1%) fail the Kolmogorov test and can be rejected. Looking closely to bad strips it appeared that most of these strips belong to a limited number of modules. A combined probability of rejection is built with the 20 worse strips and a final 2% cut on the number of bad strips is applied to reject APVs.

3.8 Cosmic rays analysis at CRAFT

After the whole preliminary calibration described so far, the Tracker has been inserted in Global Run together with all CMS detectors and with the magnetic field switched on, and it has started to collect cosmic rays data triggered by the CMS Drift Tubes.

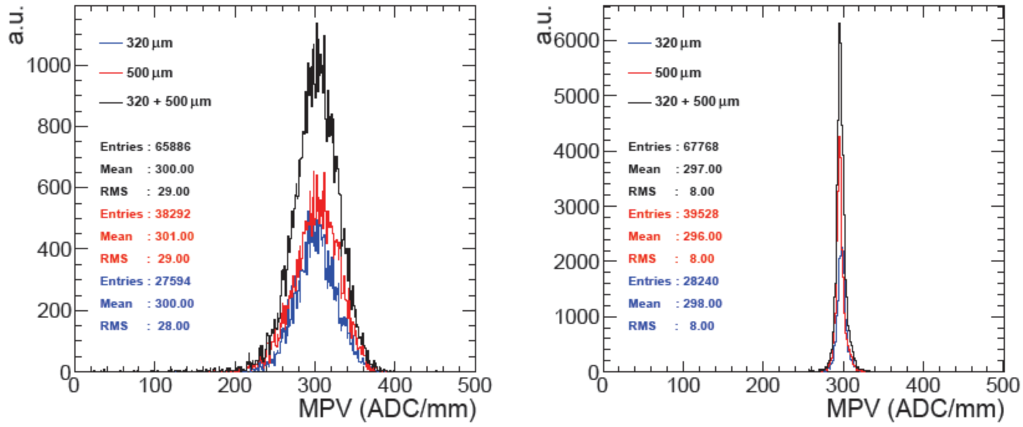
In the following the results obtained analyzing the events collected are shown.

3.8.1 Gain calibration with cosmic ray data

The charge released in the silicon by the passage of an ionizing particle is processed by the readout electronics chain, which includes the APV and AOH chips, the optical fibres and the FED. The ratio of the ADC counts output after FED digitization to the originally-released charge corresponds to the gain of the electronics chain.

Particle identification using energy loss in the silicon detectors is known to be sensitive both to the absolute calibration scale and to gain non-uniformities. It is therefore important to correct these non-uniformities and to measure precisely the conversion factor between $\mathrm{d}E/\mathrm{d}x$ and ADC counts.

The electronics gain can be made uniform throughout the tracker simply scaling the tickmark heights measured during commissioning (Sec. 2.3.1). This procedure however does not take into account non-uniformities in the silicon or between the channels within an APV, nor it will account for the other sources of possible non-uniformity. These kinds of non-uniformities can only be correctly accounted for by determining inter-calibration constants using the signals produced by particles.



(a) (b)

Figure 3.32: *Most probable value (MPV) of the cluster charge: (a) before and (b) after gain calibration.*

In order to determine the inter-calibration constants, the clusters associated with tracks were used to produce path length corrected charge distributions for each APV. The resulting distributions were fitted with a Landau curve. The most probable value of each distribution is then used to compute the inter-calibration constant by normalizing the signal to 300 ADC counts/mm - the value expected for a minimum ionizing particle with a calibration of 270 e^-/ADC count (Sec. 2.3). The distribution of the most probable values, for modules with at least 50 clusters, is shown in Fig. 3.32(a) separately for 320 μm and 500 μm thick sensors. The spread of this distribution is of the order of 10%.

The inter-calibration constants determined with this procedure were used in the final data reprocessing of the CRAFT data, resulting in a uniform response, as shown in Fig. 3.32(b).

3.8.2 The signal-to-noise ratio

The signal-to-noise ratio is considered a benchmark for the performance of the tracker. It is particularly useful, as already stated, to spot possible problems in the calibration procedures and for studying the stability of the tracker over time.

The signal-to-noise ratio distributions, corrected for the path length, are presented in Fig. 3.33 for TIB layer 1 and TOB layer 5. The distributions, obtained as output of the DQM module introduced in Sec. dqmMonitoring, are fitted with a Landau function convoluted with a Gaussian function to determine the most probable value for the signal-to-noise ratio.

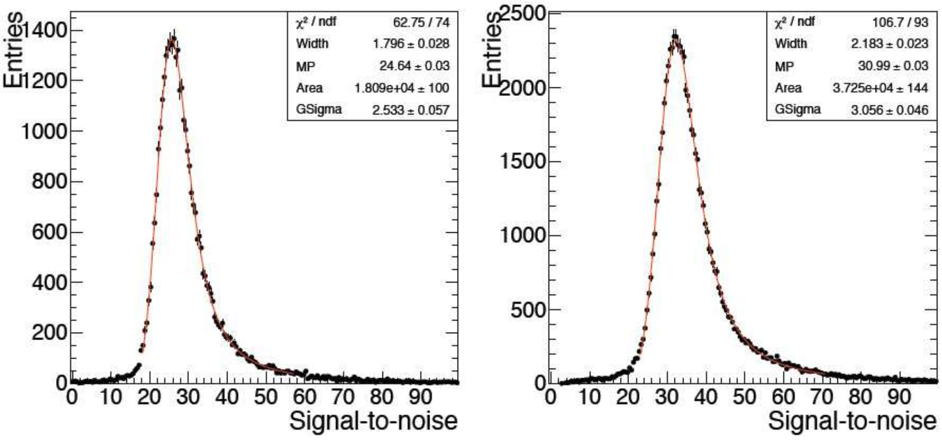


Figure 3.33: *Signal-to-noise ratio distribution of clusters associated to tracks in TIB (left) and TOB (right).*

The fit of the signal-to-noise ratio can also be performed on a run-by-run basis; Fig. 3.34 shows the most probable value of the signal-to-noise ratio as a function of run number, allowing the stability of tracker performance to be monitored over a period of time. Fig. 3.34 is divided into the three main data-taking periods defined in Table 3.6. In period A the signal-to-noise ratio value has been found to be lower than expected from previous measurements at TIF. After investigation, it was clear that the reason was that a significant fraction of the tracker was not correctly synchronized with the rest of CMS (Sec. 3.7.2.1). This has been a typical example in which an efficient monitoring tool has helped in spotting out

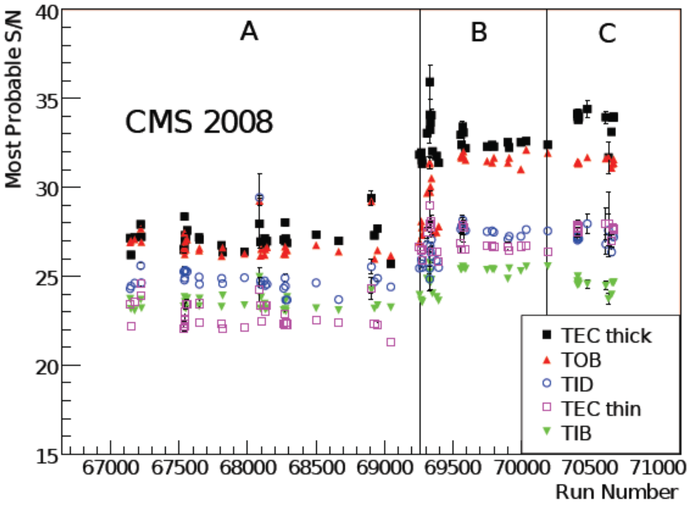


Figure 3.34: *Signal-to-noise ratio versus the run number. The error bars are the uncertainty associated with the Landau fit.*

a problem. After having tuned the synchronization, in periods B and C of data taking, the value has been found to be instead compatible with expectations and previous measurements at TIF and very stable. As expected, the signal-to-noise ratio is larger in thick sensors, TOB and TEC Rings 5-7, than in thin sensors, TIB, TID and TEC Rings 1-4. As the noise is larger in thick sensors but the deposited energy increases linearly with the thickness, thick sensors collect about 5/3 more charge. In TEC and TID, finally, the signal-to-noise ratio is even larger than in TOB and TIB, because end-cap modules have a slightly lower noise than the 300 barrel modules with the same thickness.

3.8.3 The Lorentz angle measurement

The measurement of the Lorentz angle has been introduced in Sec. 3.3.3 for the TIF data taking. The main difference here is that during Global Run the magnetic field was switched on, hence a deviation from the zero value should be seen in the fit of the cluster size versus the angle.

A profile plot of cluster size versus the tangent of the incidence angle obtained during the Global Run is shown in Fig. 3.35 for a module of TOB Layer 4.

The function $f(\theta_t)$ used to determine the minimum is again

$$f(\theta_t) = \frac{t}{P} \cdot p_1 \cdot |\tan\theta_t - p_0| + p_2 \tag{3.7}$$

where t is the detector thickness, P is the pitch and p_{0-2} are the fitted parameters. The most important parameter, p_0, is the estimate of $\tan\theta_L$,

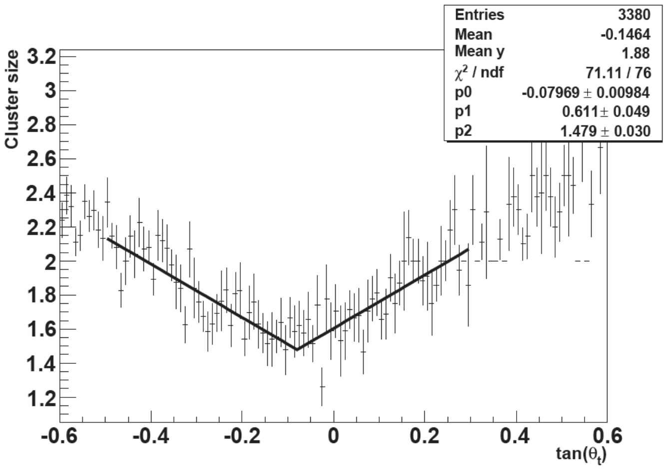


Figure 3.35: *Cluster size versus incident angle in one module of TOB Layer 4 at 3.8 magnetic field.*

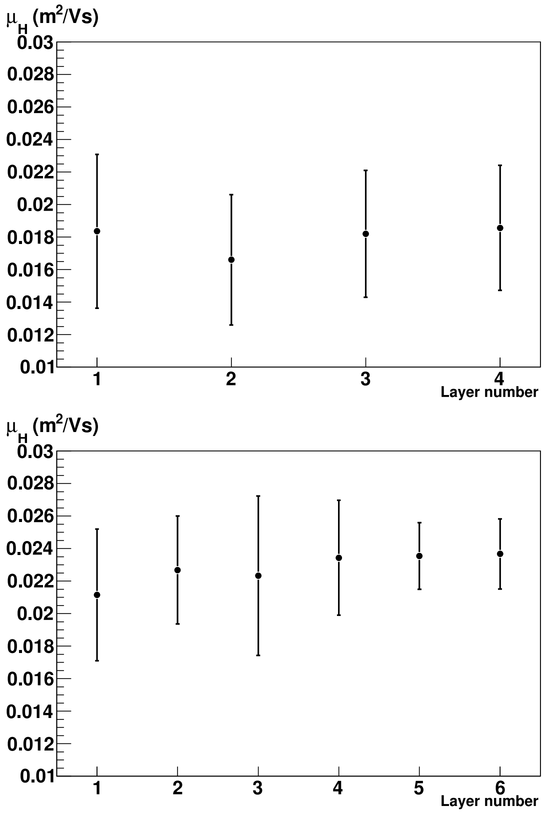


Figure 3.36: *Measured values of* μ_H *for* $B = 3.8$ *T in TIB (top) and TOB (bottom).*

p_1 is the slope normalized to the ratio of thickness over pitch and p_2 is the average cluster size in the minimum.

Given that the Hall-mobility is defined as $\mu_H = \tan\theta_L/B$, where B

is the magnetic field in Tesla, the result of the fitting procedure is, in effect, a measurement of μ_H. The mobility is measured for each individual module, so the value for each layer can be determined simply as an average of the mobilities measured for each module in the layer. The results for the barrel are shown in Fig. 3.36, where the uncertainties quoted are the variances of the measured values in the layer. The results of the layer by layer measurements are clearly consistent in TIB and in TOB, but a small difference is observed between the two subdetectors because of the different thickness.

The measured values are used during the track reconstruction, and have been confirmed by the tracker alignment validation [75].

3.8.4 Hit efficiency

A single run from the CRAFT dataset has been used in order to assure that the number of excluded modules did not change. A very long run was chosen to dispose of a sufficient track statistics. There were between 16 400 and 104 800 tracks per barrel layer and between 1 700 and 6 500 per end cap layer. The analysis was limited to events with only one track, which was required to have a minimum of eight hits and no more than four missing hits. To ensure that the muon has actually passed through the module under study, the location of the extrapolation of the track trajectory on the module surface was required to lie at more than a certain distance from the sensor edge. That distance was defined to be five times the position uncertainty of the extrapolated point.

Partition	TEC+	TEC-	TOB	TIB/TID	SST
Modules in system	3200	3200	5208	3540	15148
Functional modules	3189	3198	5196	3487	15070
Fraction	(99.7%)	(99.9%)	(99.8%)	(98.5%)	(99.5%)
Module used	3175	3144	5106	3422	14847
Fraction	(99.2%)	(98.3%)	(98.1%)	(96.7%)	(98.0%)

Table 3.8: *Statistics of functional modules after the commissioning procedures.*

The efficiency results per tracker layer are shown in Fig. 3.37. These measurements, which include all tracker modules, are compatible with the expected overall percentage of excluded modules reported in Table 3.8. If the modules that were excluded because of known problems were ignored in the efficiency calculation, the resulting efficiency would be greater than 99% for most layers. No more than about 0.1% of the inefficiency arises from isolated dead strips, which are in total 21 574 and have not been taken into account in the efficiency calculation. The rest is attributed to modules that were problematic only for a short period of time and were therefore not identified by the other procedures described

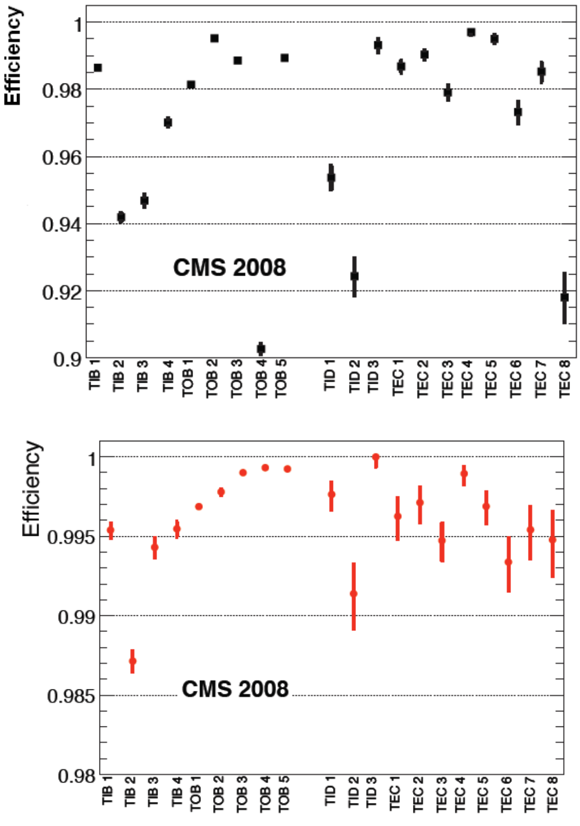


Figure 3.37: *Average module hit efficiency per layer/disks, including (top) or excluding (down) modules disconnected or otherwise excluded. The efficiency cannot be measured in the outermost layers of TOB (layer 6) or TEC (layer 9) without modifying the track reconstruction algorithm, because the track reconstruction requires the presence of a hit in the outermost layer or disk, depending on the track trajectory.*

in this paper. Subsequent improvements, such as detailed logging of modules affected by temporary power supply problems (HV trips etc.), will remove this source of inefficiency from the monitoring for future data-taking.

3.8.5 Tracking performance

The two main algorithms used to reconstruct tracks from cosmic muons in CRAFT data are the Combinatorial Track Finder (CTF) and the Cosmic Track Finder (CosmicTF).

There have been two significant changes in the Combinatorial Track Finder since its first use in the Slice Test, both relating to the seed finding phase. The Sector Test at TIF was performed without the presence of a magnetic field and with only limited angular coverage. Now that the full tracker is available, seed finding in the barrel uses TOB layers only.

In the end caps, hits in adjacent disks are used to form hit pairs. The presence of the 3.8 T magnetic field means that for hit-triplet seeds, the curvature of the helix yields an initial estimate of the momentum. For hit pairs seeds, an initial estimate of 2 GeV/c is used, which corresponds to the most probable value.

The results presented here are based on data taken in Period B. This ensures that data are unaffected by synchronization problem discussed in Sec. 3.6.1 and that the data taking conditions are approximately uniform. Period C is not used because it corresponds to data taken with no magnetic field.

The number of tracks reconstructed by the two algorithms in the data from Period B, without applying any additional track quality criteria, except those applied during the track reconstruction itself, are 2 196 949 using the Combinatorial Track Finder and 2 785 396 using the Cosmic Track Finder.

The number of reconstructed tracks per event is shown in Fig. 3.38, where also a comparison with Monte Carlo simulation is shown. The large number of events without reconstructed tracks is mainly due to muons outside the fiducial volume for which less than five hits are reconstructed in the tracker.

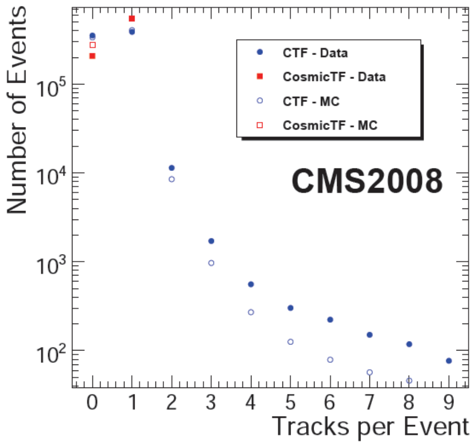


Figure 3.38: *Distribution of the number of reconstructed tracks per event for cosmic muon tracks reconstructed with the two different algorithms. For each algorithm, the total number of simulated Monte Carlo tracks are normalized to the number of observed tracks.*

Fig. 3.39 shows the distribution of χ^2 and transverse momentum compared between the two algorithm and between data and Monte Carlo simulation.

It can be seen that reasonable agreement is found between data and Monte Carlo simulation, although there are some discrepancies that are thought to be due to the reconstruction of showers by the track recon-

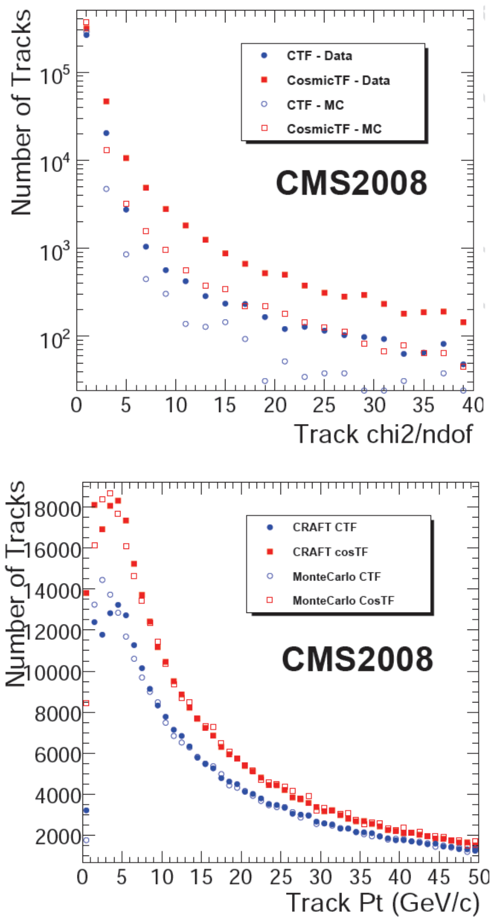


Figure 3.39: *Distribution of* χ^2/ndf *(top) and transverse momentum (bottom) for cosmic muon tracks reconstructed with two different algorithms for single track events. Note that for the* χ^2/ndf *a log-scale is used for y-axis. For each algorithm, the total number of simulated Monte Carlo tracks is normalized to the number of observed track.*

struction algorithms. The CTF is capable of reconstructing more than one track per event, but as it has not been optimized to reconstruct showers, multi-track events tend to contain a number of fake or badly reconstructed tracks. These are mostly low momentum tracks with a small number of hits and large χ^2 values, and the fake rate is estimated to be around 1%. For this reason, only single track events are used for the analysis here presented.

By design the Cosmic Track Finder reconstructs only one track. The difference between the number of tracks reconstructed by the two algorithms is mainly due to the minimum number of hits required during the pattern recognition phase: in the CTF a minimum of five hits are required, while only four hits are required in the case of the CosmicTF.

3.8.6 Track reconstruction efficiency

The track reconstruction efficiency for the two algorithms described above has been measured using two different methods. First, the efficiency was measured by searching for a reconstructed track and matching it to a muon reconstructed only in the muon chambers. In the second method, the efficiency was measured using data just from the tracker, by reconstructing tracks independently in the upper and lower hemispheres of the tracker.

In the first method, the track reconstruction efficiency is measured with respect to muons reconstructed using information from the muon chambers, and required to point within the geometrical acceptance of the tracker. This ensures that the muons have been identified independently of the tracker.

The muons are first reconstructed by the muon chambers, combining segments of muon tracks reconstructed in the top and bottom hemispheres of the muon detectors in a global fit. These reference muons are required to have at least 52 hits in the muon chambers, which corresponds to having hits in at least five Drift Tube chambers. Combining segments from the two hemispheres removes muons which are absorbed by the CMS steel yoke before reaching the tracker. It also improves the track direction reconstruction, which is needed for the propagation through the detector. To ensure that the muon does indeed cross the tracker and traverses enough layers to be reconstructed by the algorithms, its trajectory is propagated into the tracker and is required to point inside a cylinder surrounding TOB layer 4. The efficiency is then measured by searching for the corresponding reconstructed track in the tracker.

The efficiency is estimated with respect to reference muons with a topology similar to the one expected in proton-proton collisions. This is achieved by requiring that the point of closest approach of the extrapolated muon to the centre of the detector is less than 30 cm in both the transverse and longitudinal directions. The absolute value of the pseudorapidity, $|\eta|$, is required to be less than 1 and the azimuthal angle is required to be in the range $0.5 < |\phi| < 2.5$, effectively restricting the tracks to the barrel. These cuts also ensure that the tracks cross most of the layers of the tracker and cross most modules perpendicularly. The efficiencies measured in the data and in the Monte Carlo simulation are compared in Fig. 3.40 and summarized in Table 3.9.

The efficiencies are higher than 99% for both data and Monte Carlo simulation for the two tracking algorithms. The difference between data and Monte Carlo observed for values of transverse momentum around 20 GeV/c for the Cosmic Track Finder, while statistically significant, is small and has not been pursued further, since this algorithm will not be used in proton-proton collisions. The overall differences between data and Monte Carlo simulation are found to be smaller than 0.5%.

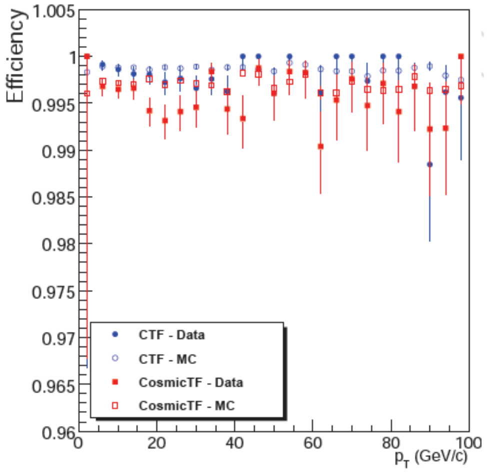


Figure 3.40: *Track reconstruction efficiency as a function of the measured transverse momentum of the reference track, as measured with the track-muon matching method.*

	Data (CTF)	MC (CTF)	Data (CosmicTF)	MC (CosmicTF)
Efficiency	99.78 ± 0.02	99.88 ± 0.00	99.47 ± 0.04	99.72 ± 0.01

Table 3.9: *Track reconstruction efficiencies for the two algorithms in Data and in Monte Carlo simulation, measured with the muon-matching method.*

In the second method, the efficiency is measured using data from the tracker; no muon chamber information is included. This removes any possible bias arising from the reconstruction and the selection of the muons or from matching muon chamber and tracker tracks. This method had been already used during Sector Test, on a smaller fraction of the tracker. As cosmic muons pass through the tracker from top to bottom, the tracker was divided into two hemispheres along the $y = 0$ horizontal plane for this study. The tracks were reconstructed independently in the two hemispheres. Tracks reconstructed in the upper hemisphere are referred to as top tracks and those reconstructed in the lower hemisphere as bottom tracks. In order to measure the efficiency in one of the hemispheres, reference tracks in the other hemisphere are used. The hemisphere under study is then searched for matching tracks. Two efficiency measurements can then be performed: $\epsilon(\mathrm{T|B})$, where, given a bottom track, a matching top track is sought, and vice versa ($\epsilon(\mathrm{B|T})$). The matching is performed by requiring that the two opposite-half tracks have pseudorapidities that satisfy $|\Delta\eta| < 0.5$. Only events containing a single track with a topology similar to the one expected in proton-proton collisions are analysed and the same track requirements that were applied for the first method were used. In order to reconstruct the two track legs independently, only seeds with hits in the top or bottom hemisphere are selected and, before the final track fit, the hits in the other hemisphere are removed from the track. After track segment reconstruction, a track is only retained for further analysis if it contains at least 7 hits and its χ^2 satisfies the requirement $\chi^2/\mathrm{ndf} < 10$. Furthermore, to ensure that a matching track can be reconstructed, the extrapolation of the reference track into the other hemisphere is required to cross at least five layers.

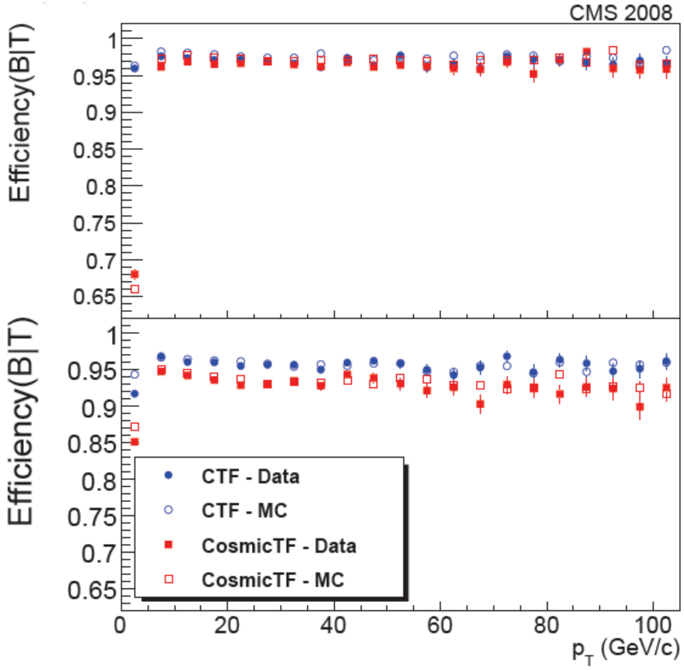


Figure 3.41: *Track reconstruction efficiency as a function of the measured transverse momentum of the reference track, as measured with top-bottom comparison method.*

	Data (CTF)	MC (CTF)	Data (CosmicTF)	MC (CosmicTF)
ϵ(B|T)(%)	97.03±0.07	97.56±0.04	94.01±0.10	93.41±0.06
ϵ(T|B)(%)	95.6±10.08	95.79±0.05	92.65±0.11	93.19±0.07

Table 3.10: *Overall track reconstruction efficiency measured with the top/bottom comparison method.*

The efficiencies measured using this method are summarized in Table 3.10 and shown in Fig. 3.41. The difference seen for low momentum tracks for the Cosmic Track Finder is small, and has not been pursued further. The lower efficiency for top tracks is primarily caused by a large inactive area in the upper half of TOB layer 4, which would otherwise be used to build track seeds. This will not be an issue for the track reconstruction that will be used in proton-proton collisions as in this case, tracks are seeded principally in the pixel detector with the tracking then proceeding towards the outer layers of the tracker. The efficiencies measured in the Monte Carlo simulation are consistent with those measured in the data to within 1%, provided that the known detector inefficiencies are accounted for in the simulation.

3.8.6.1 Inside-out tracking method

Finally, in order to evaluate the tracking efficiency of the algorithm that will be used during proton-proton collisions, i.e. the Combinatorial Track Finder, the track reconstruction process has been repeated starting from the centre of the tracker and proceeding inside-out, emulating the reconstruction in the LHC collisions. The default Combinatorial Track Finder is optimized to reconstruct tracks that originate near the interaction point. By contrast, very few cosmic muons pass through this region. In order to take this into account, only tracks for which the point of closest approach to the centre of the detector is less than 4 cm in the transverse direction and 25 cm in the longitudinal direction are used. This requirement assures that the tracks effectively cross the three barrel layers of the pixel detector.

The tracks are reconstructed from a seed made with hit pairs from any combination of the innermost three layers of the tracker; the nominal beam spot is used as an additional constraint in the transverse plane to provide the initial estimate of the track parameters. This is a legitimate approximation as long as the transverse impact parameter of the tracks is much smaller than the radius of the innermost detector layer used. Hits in the silicon pixel detector are not used in this analysis in the seed finding phase, as this imposes too strong a constraint on the tracks to come from the nominal beam spot. They are, however, identified in the pattern recognition phase and added to the track.

The reconstruction efficiency is estimated with respect to a reference track in one hemisphere of the tracker. A compatible seed and track is sought in the other hemisphere within a cone of radius $\Delta\mathrm{R} < 1.0$ (where $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) opposite to the reference track. The cone size is kept very large compared to the angular resolution so that the matching procedure cannot bias the efficiency measurements. To avoid multi-track events, a track is not used as a reference if there is another track in the same hemisphere or within the matching cone. Fake tracks created by noisy hits are rejected by requiring that the reference tracks have at least 10 hits. The efficiency measured using this method is shown in Fig. 3.42 and in Table 3.11 as a function of the transverse momentum of the reference tracks, for the Monte Carlo simulation and for cosmic data.

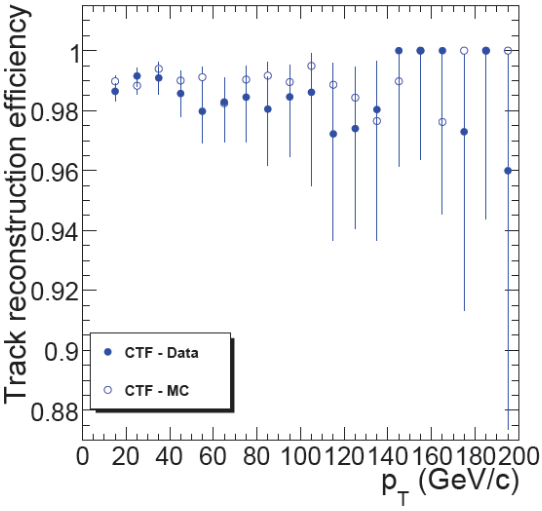


Figure 3.42: *Track reconstruction efficiency as a function of the measured transverse momentum of the reference track for inside-out tracking method.*

Inefficiencies affecting only a few detector channels have not been taken into account when calculating the overall efficiency from the data, while the efficiency calculated using the Monte Carlo simulation assume all detector channels working properly. The efficiency measured in the Monte Carlo simulation matches the one measured from data within 1%, providing known detector problems are taken into account.

	Data	MC
Track reconstruction efficiency (%)	98.96±0.13	98.94±0.09

Table 3.11: *Reconstruction efficiency of the inside-out tracking method.*

3.8.7 Track parameter resolution

The track reconstruction can be further validated using the CRAFT data sample by splitting the tracks into two separate parts. A measure of the resolution of the track parameters can be determined by comparing the two legs of the split tracks. To perform this study, tracks are split at the point of closest approach to the nominal beam-line. The top and bottom legs are treated as two independent tracks and re-fitted accordingly. The track parameters are then propagated to their respective points of closest approach to the beam-line. This method has been tested using Monte Carlo simulation and is found to work well.

For the purposes of this study, only events in which the Combinatorial Track Finder reconstructed a single track whose point of closest approach to the beam-line is inside the volume of the pixel barrel are considered. The transverse momentum of the track must be larger than 4 GeV/c and its χ^2 must satisfy the requirement $\chi^2/\mathrm{ndf} < 100$. In addition, the track must contain a minimum of 10 hits, with at least two hits being on double-sided strip modules. There must also be six hits in the pixel barrel subsystem. After splitting, each track segment is required to have at least six hits, three of which must be in the pixel barrel.

Using these selected tracks, the residuals defined as $\delta x = (x_1 - x_2)/\sqrt{2}$ and the standardized residuals (or pulls) $\widetilde{\delta x} = (x_1 - x_2)/\sqrt{\sigma_{x1}^2 + \sigma_{x2}^2}$ are calculated. The indexes 1 and 2 stand for the two track legs, and the factor $\sqrt{2}$ is needed to account for the fact that the two legs are statistically independent.

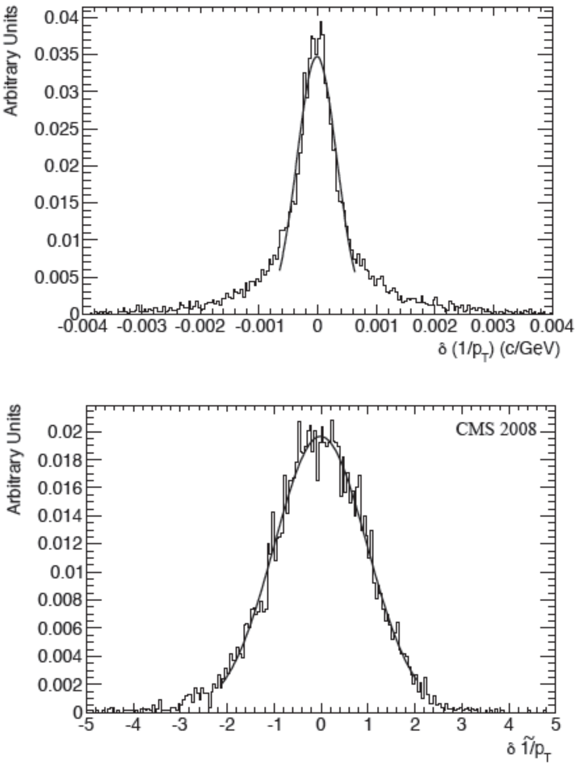


Figure 3.43: *Residual distribution (top) and pull distribution (bottom) of the inverse transverse momentum* $1/p_T$.

The distribution of the residuals and the pulls of the inverse transverse momentum are shown in Fig. 3.43, the same distributions for the azimuthal (ϕ) and polar (θ) angles are reported in Appendix Fig. 7.

In Table 3.12 the mean and the standard deviation (referred to as the *resolution*) of a Gaussian fitted to the peak of the distributions are given. The same quantities are used to characterize the pull distributions. In

	Residual distributions		Pull distributions	
	Std. Dev.	Mean	Std. Dev.	Mean
p_T (GeVc^{-1})	0.083	0.000	0.99	0.01
$1/p_T$ (GeV^{-1}c)	0.00035	0.00003	0.99	-0.01
ϕ (mrad)	0.19	0.001	1.08	-0.02
θ (mrad)	0.40	0.003	0.93	-0.01
I.P. $\perp$ (μm)	22	0.30	1.22	0.00
I.P $\|$(μm)	39	0.28	0.94	-0.01

Table 3.12: *Standard deviation and mean of the residual and pull distributions of the track parameters (I.P. = Interaction Point). The units indicated pertain only to the residual distributions.*

this case, the standard deviations of the fitted Gaussians are taken as the pull values. It can be seen that the transverse momentum shows significant tails. This is due to the fact that tracks within a wide range of momenta are used and that the resolution degrades with increased momentum.

3.9 Tracker alignment with cosmic rays data

Starting from the experience gained with the alignment of about 15% of the CMS silicon strip tracker during the stand-alone commissioning at TIF, significant improvements have been introduced in the alignment test during CRAFT. Inclusion of the pixel detector and of the full strip tracker made the first alignment of the whole CMS tracker possible. The magnetic field allowed an estimate of the multiple scattering effects on a track-by-track basis. Two complementary statistical alignment algorithms, a global and a local iterative method, described in Sec. 2.6.3, were applied to data. Although both algorithms had been used previously during standalone commissioning, further improvements have been introduced, such as inclusion of survey data together with tracks. Each of the two alignment algorithms was used to obtain module positions independently and a comparison of the results between the two complementary approaches was part of the validation of the tracker alignment procedure. After verifying that the two methods yielded consistent results, the final results were obtained by applying the two algorithms in sequence in order to take advantage of their complementary strengths.

The alignment analysis was performed after the proper calibration of both the strip and the pixel detectors. In particular, an appropriate Lorentz angle calibration was applied. A Lorentz angle miscalibration would result in a systematic shift of the effective charge collection position, in the same u-direction for all the modules with the same relative orientation between the solenoid and the drift field. Since this would appear as an effective shift of the modules, it would not be evident if corrected for by the alignment. Therefore, for validation purpose only, alignment results were checked with the data collected with 0 T and with 3.8 T solenoid fields. The precision achieved in the alignment analysis is such that there should be clear evidence for a Lorentz angle miscalibration with an effective shift of only few microns. In fact, good agreement with the Lorentz angle calibration was found.

Several approaches were employed to validate the alignment results. The low-level quantities used in the χ^2 minimization, such as residuals and the χ^2/ndf of the tracks, were monitored.

The track χ^2/ndf distribution is shown in Fig. 3.44. The hit residuals for TIB and TOB are shown in Fig. 3.45

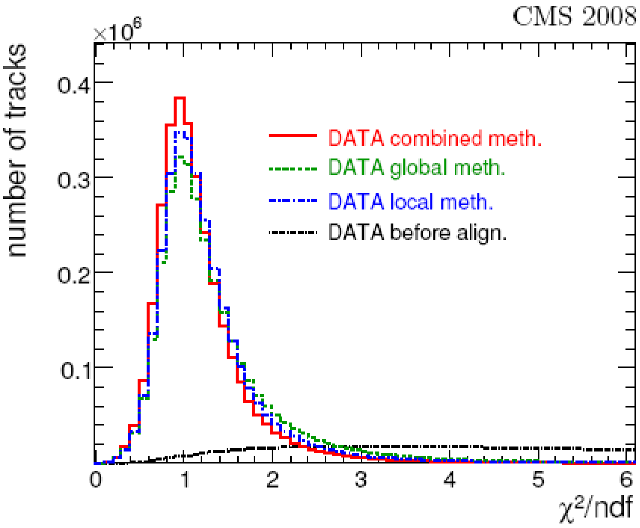


Figure 3.44: *Distribution of the* χ^2*/ndf of the tracks before alignment (black), and after alignment with the local (blue), global (green) and combined (red) methods.*

The hit residual width is dominated by two effects other than alignment: track extrapolation uncertainties due to multiple scattering and hit position reconstruction uncertainties. Both of these effects are random, while misalignment leads to systematic shifts of the residuals. For this reason, the distribution of the median of the residuals (DMR) is taken as the most appropriate measure of alignment. Median distributions are shown in Fig. 3.46.

The results show that the application of the two method combined give an excellent result, extremely close to the ideal Monte Carlo alignment scenario in TIB and TOB. In TID and TEC the result is as well very good considering the difficulty to align endcap regions with cosmic rays. The precision achieved on module position measurement in the

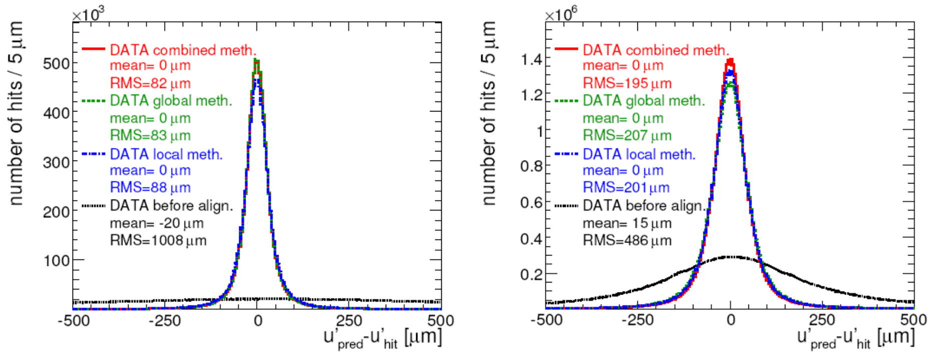


Figure 3.45: *Track residuals for TIB (left) and TOB (right). The four lines correspond to positions before alignment (black) and alignment with the global (green), local (blue) and combined methods (red).*

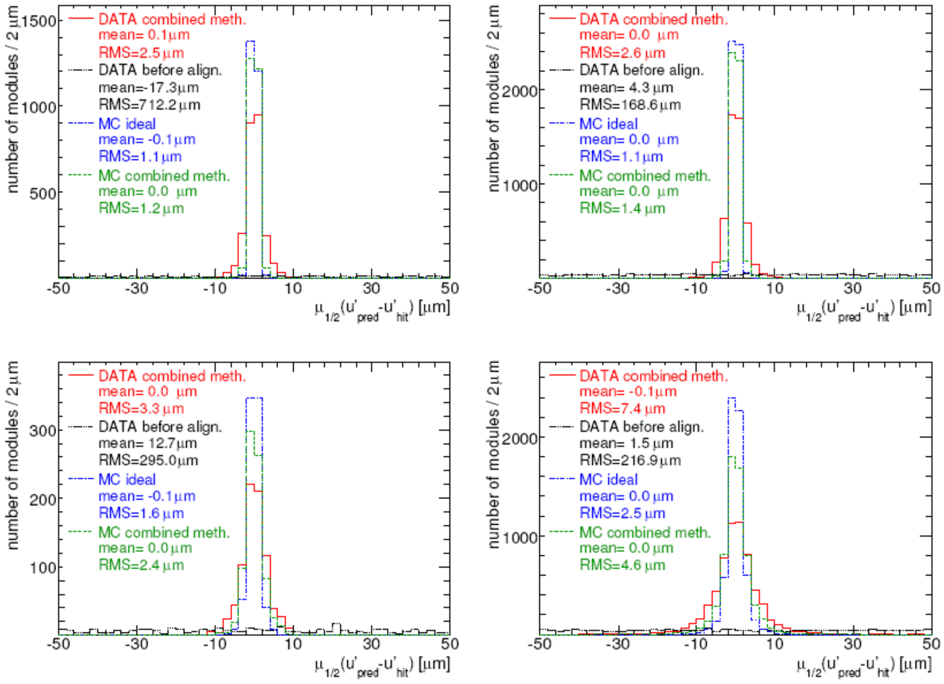


Figure 3.46: *Distribution of the median of residuals for modules with more than 30 hits, shown for TIB (top left), TOB (top right), TID (bottom left) and TEC (bottom right). The distributions are shown before alignment (black) and alignment with and after the alitnment with global (green), local (blue) and combined methods (red).*

most sensitive coordinate results to be about 2.5 μm in the barrel and ranging from 3.3 to 7.4 μm in the endcaps, respectively. This level of precision is much higher than the one obtained during the Sector Test and much closer to the alignment precision expected with collision data.

3.10 Summary on CRAFT

The Cosmic Run at Four Tesla has been an important experience for the commissioning of the tracker.

The control and readout system were successfully commissioned, synchronized to the Level-1 trigger and operated in global runs with all the other subdetectors of the CMS experiment, for the first time before collisions. The total number of modules used corresponds to 98% of the total system.

About 13 million events with a muon in the tracker were collected with the magnet on. The hit and track reconstruction have shown excellent performances and the CTF default algorithm that will be used in proton-proton collisions was tested successfully. The signal-to-noise performance exceeded the detector design specifications, being in the range 15-30 for thin modules and 31-36 for thick ones. The efficiency of the hit reconstruction is above 99.5%. In addition, with the collected statistics, a measurement of the Lorentz Angle has been possible.

The track reconstruction efficiency has been measured with two different methods: one using muons reconstructed in the tracker and in the muon chambers and one using only data from the tracker. The reconstruction efficiency in data was found to be high and well described by the Monte Carlo simulation. For tracks which are close to the centre of the detector and which have a direction close to the vertical axis, the track reconstruction efficiency was found to be higher than 99%. The resolution on hit position and track parameters was also consistent with expectations from Monte Carlo simulation.

The alignment has shown a strong improvement with respect to the precision achieved during Sector Test, reaching precision on module position measurement in the most sensitive coordinate of about 3 μm in the barrel and less than 8 μm in the endcaps combining the two alignment methods. This alignment is compatible with the maximum statistics precision that can be achieved using cosmic tracks, which do not come from the interaction point.

The CRAFT was an important milestone towards final commissioning with colliding beam data and it showed that design performance can be achieved.

The results obtained during CRAFT data taking will be used in next Chapter to simulate a realistic detector condition in early data taking.

Chapter 4

Z boson physics in early data for the momentum resolution measurement and the Monte Carlo simulation tuning

In order to enable the CMS experiment to perform precise measurements, it is mandatory to carefully calibrate its detector subsystems with first collisions data.

Among the measurements provided by CMS, track parameters have a paramount importance. Transverse momentum in particular is the most subjected to details of the detector integration and operation, being highly sensitive to the precise alignment of the silicon sensors of the tracker and of the muon chambers, to the material budget in the detector, to the detailed map of the magnetic field inside and outside the solenoid volume, and in general to the performance of tracking detectors. The inner silicon tracker in particular has a leading role in the momentum resolution of muons up to 200 GeV/c, therefore its calibration and alignment conditions will be a pivotal requirement during the real data taking in order to achieve the best possible resolution in this range.

The calibration of track momentum and the investigation of possible residual systematic effects introduced by the above sources and not completely addressed, is greatly eased by the availability of reconstructed resonance decays, typically, although not exclusively, those involving two-body decays of neutral particles (i.e., Υ, J/Ψ, Z,...). Z boson is a perfect candidate to perform this kind of measurement, given the high production cross-section and the extremely low background. It is produced via Drell-Yan process and its theoretical cross section is calculated at next-to-next-to-leading order. A total cross section of 2.5095 ± 0.0241 nb is expected at 10 TeV proton-proton center of mass energy. This is a relevant cross section, which will permit to collect a large amount of Z events already in the early phase of data taking.

Although the muon pairs originated from the Z decays have momenta in the range from a few tens of GeV/c to less than a hundred GeV/c, and are thus hardly useful for a direct check of the momentum measurement in the TeV range, they do constitute an invaluable tool to spot deficiencies in our Monte Carlo description of the detector, becoming beneficial to tracks of any momentum. A well-calibrated momentum scale is mandatory for precise measurements of the top quark and W boson masses, for B-hadron spectroscopy, and for a number of other first-class measurements that CMS is expected to deliver. If handled properly, the large signals of di-muon resonances like Z boson allow also a precise determination of track momentum resolution, whose knowledge is as important as the one of the momentum scale in several physics measurements, and whose modeling in the simulations is very hard.

In the following it will be shown how a determination of both momentum scale and resolution can be extracted with the same methodology from suitable sets of data, and how from the measurement of the resolution on a real data sample it is possible to tune the resolution in the Monte Carlo simulation, already with the first 10 pb^{-1} integrated luminosity.

This Chapter is organized as follows: an introduction to the algorithm used to measure the muon resolution from data and to correct the muon momentum scale is given in Sec. 4.1. A description of the generation and the reconstruction of the samples used for the analysis follows. In Sec. 4.3 the results of the scale and resolution measurement on different Z samples are reported. Finally in Sec. 4.4 the description of the method to tune the Monte Carlo resolution on muon momenta using first collision real data is presented, together with the results.

4.1 Method description

Following the motivation just outlined, an algorithm has been developed, capable of extracting all the information available from the combination of the reconstructed kinematics of the muon pair with the supposed knowledge of the parent particle species. The reference quantity on an event-by-event basis to the possible biases on the reconstructed muon track parameters is, of course, the parent's invariant mass. This is not a per-track variable, so a probabilistic approach is necessary in order to relate the difference between expected and observed mass with a hypothetical bias on the measured parameters of either or both daughter tracks.

The connection between scale and resolution in the determination of measurement biases is obvious: a potential bias in the momentum scale can be meaningfully inferred from the measured di-muon mass on an event-by-event basis only if a value is assumed for the uncertainty in the quadri-momentum of the daughter tracks, such that an estimate of mass resolution is possible. The information on the scale is hidden by the finite mass resolution, but can be fully recovered by a probabilistic approach.

A multi-parameter likelihood fit is the best approach to the solution of this problem. Once one defines a set of functions describing the dependence of the biases (offsets from the true values) on track kinematics and on the measurement resolutions, the best estimate of the functions parameters can be determined from a likelihood minimization, provided that a sufficient set of homogeneous data is used. The problem thus simply becomes finding the functions which describe most accurately those biases and resolutions. Their features have a direct correspondence with the possible shortcomings of the description of the detector material, of the alignment of the sensitive devices, of the magnetic field map, and of tracking reconstruction: by experiencing with different simulated situations, it is possible to build a library of functions addressing those potential shortcomings. Once real data become available, it will be possible not only to correct the observed biases, but also to judge what are their sources, enabling in some cases their removal.

The details of the method are the following. In proton-proton collisions, Z bosons are predominantly produced via quark-antiquark annihilation. The initial state partons “sample” the relativistic Breit-Wigner shape of Z production cross-section

$$\hat{\sigma}(m) = \frac{\Gamma^2 M_{ref}^2}{[(x^2 - M_{ref}^2)^2 + \Gamma^2 M_{ref}^2]} \tag{4.1}$$

according to their parton distribution functions (PDF) $f_i(x, m^2)$, thereby producing a modified lineshape, as dictated by the following factorization integral (in leading-order approximation):

$$\sigma(m) = \Sigma_{i=udscb} \int dx_1 \int dx_2 \, k_i^2 f_i(x_1, m^2) f_{\bar{i}}(x_2, m^2) \, \hat{\sigma}(m) \delta(m^2 - x_1 x_2 s) \tag{4.2}$$

where $s = 100$ TeV^2, k_i is the appropriate weak factor, and where parton distribution functions are evolved to the actual scale m^2 at which Z production occurs.

The muons produced in the decay of the Z boson then cross the detector material and their trajectories are reconstructed starting from the hits in the silicon layers and in the muon detectors.

With measured muon kinematics, a reconstruction of the mass of the parent Z boson results in a resolution-smeared view of the modified Breit-Wigner lineshape $\hat{\sigma}(m)$ describing the produced Z bosons. A bias will manifest as a shift of the most probable value from the one of the produced Z boson, while the effect of resolution will result in a widening of the distribution[1]. If m is assumed to be the measured mass of the di-muon pair, and s the resolution on the measured mass, derived from a suitable model parametrizing the momentum resolution as a function of muon kinematics, the probability can be defined, modulo a normalization factor, as:

$$P(m,s) \;=\; \int dx \frac{\sigma(x)}{s}\, e^{-\frac{(x-m)^2}{2s^2}} \tag{4.3}$$

In order to construct a likelihood function to be minimized for this purpose, the observed reconstructed mass m and the expected mass resolution s must therefore be written as a function of the parameters which describe the momentum scale correction function and the resolution on the measured quantities. For this purpose muon momenta are expressed as a function of the measured variables $(p_T, \phi, \cot\theta)$:

$$\begin{cases} E = \sqrt{p_T^2(1+\cot^2\theta)+m_\mu^2}; \\ P_x = p_T\cos\phi; \\ P_y = p_T\sin\phi; \\ P_z = p_T\cot\theta. \end{cases} \tag{4.4}$$

The parametrization of scale correction can be thus written as

$$p_T^{corr} = F(\vec{x};\vec{\alpha})\, p_T, \tag{4.5}$$

such that F is understood to be close to unity. In the above formula $\vec{x}$ represents all the observable quantities on which the momentum scale has been found to depend; these of course may include the five track helix parameters of each muon, as well as other continuous variables (instantaneous luminosity, track isolation, etc.), and discrete ones (number and pattern of hits of the track, kind of track considered). The scale correction function parameters $\vec{\alpha}$ thus enter in the likelihood indirectly through the dependence of the reconstructed mass on the corrected muon transverse momenta, not affecting the angular variables:

$$m \;=\; m(P_{T,1}^{corr}, \phi_1, \cot\theta_1;\; P_{T,2}^{corr}, \phi_2, \cot\theta_2). \tag{4.6}$$

The mass resolution of the resonance, s, can be formally written in the chosen basis as

$$s = \sqrt{(\frac{\partial m}{\partial p_T})^2\sigma_{p_T}^2 + (\frac{\partial m}{\partial \phi})^2\sigma_\phi^2 + (\frac{\partial m}{\partial \cot\theta})^2\sigma_{\cot\theta}^2}, \tag{4.7}$$

and it thus depends on three additional functions, modeling the resolution in track transverse momentum, azimuth, and $\cot\theta$:

---

[1] A small downward shift of the most probable value is also expected from a significant resolution smearing, due to the asymmetric nature of the convolution between Breit-Wigner and proton PDF functions.

$$\sigma_{p_T} = G_1(\vec{x};\vec{\beta}); \sigma_\phi = G_2(\vec{x};\vec{\gamma}); \sigma_{\cot\theta} = G_3(\vec{x};\vec{\delta}). \tag{4.8}$$

Suitable forms of G_1, G_2, G_3, and initial values for their parameters $\vec{\beta}$, $\vec{\gamma}$, and $\vec{\delta}$ can be pre-determined with preliminary studies on simulated samples addressing different features.

One can finally write, for the likelihood definition, a sum of logarithms of probabilities for the N_{ev} di-muon events in the considered dataset:

$$-\ln L = -\Sigma_{k=1}^{N_{ev}} \ln(P(m_k, s_k)), \tag{4.9}$$

where m_k and s_k are understood to be computed with the measured characteristics of event k and to be functions of $\vec{\alpha}$, $\vec{\beta}$, $\vec{\gamma}$, $\vec{\delta}$. A minimization of $-\ln L$ with respect to those parameters allows the determination of scale and resolution.

The probability density functions that determine how likely a reconstructed mass value is as a function of the expected mass resolution of the di-muon pair have been evaluated previously during accurate studies using Z samples generated with `HORACE` [87] event generator. `HORACE` is a Monte Carlo event generator for Drell-Yan $pp \rightarrow \ell\nu X, \ell^+\ell^- X$ processes which includes the exact 1-loop electroweak radiative corrections matched with a QED Parton Shower to take into account also higher-order QED leading effects.

For Z boson, whose lineshape exhibits a subtle rapidity dependence due to proton PDF effects, the probability densities are a weak function of the absolute di-muon rapidity. The probability density functions are thus evaluated in rapidity bins from 0 to 2.4 in 0.1-wide rapidity intervals and stored in 24 two-dimensional surfaces, in a 1000×1000 grid on the mass-resolution plane. An example of these two-dimensional surfaces is shown in Fig. 4.1.

Given the measured muon kinematics, the algorithm computes the probability of observing a reconstructed mass m given the expected mass resolution s by interpolating the discrete grid of probability measurements to obtain a value for any (m, s) pair within their allowed support. Plugging the value of the probabilities for each event into eq. 4.9 and minimizing the likelihood, the algorithms gives as result the set of parameters of the scale and resolution ansatz functions given as input.

Of course, given the typically large number of parameters involved, a fitting strategy must be designed to obtain a stable minimum by fixing and then releasing parameters in a suitable order. The algorithm allows the maximum flexibility of configuring the number of iterations, the parameters to be released at each iteration, the order by which the release has to occur, initial parameter values, etcetera.

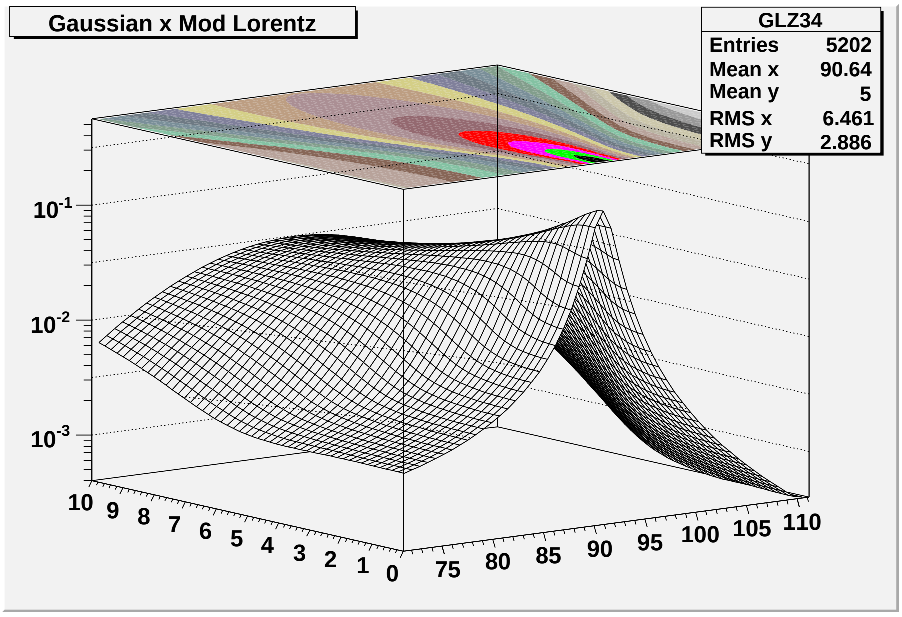


Figure 4.1: *Probability density function for reconstructed Z bosons as a function of the measured di-muon mass m (on the x axis) and the expected mass resolution s (on the y axis). The calculation of the Z boson lineshape convoluted with CTEQ parton distribution functions has been performed with the* `HORACE` *Monte Carlo. The distribution obtained for Z rapidity in the range* $1.4 < |\eta| < 1.5$ *is shown.*

4.2 Z boson samples generation and reconstruction

The algorithm has been tested on samples of about 5000 Z $\to \mu\mu$ reconstructed events, corresponding to ~10 pb^{-1} of data. The events have been produced with `SHERPA` [86], a Monte Carlo event generator for the *Simulation of High-Energy Reactions of PArticles* in a large variety of collisions, including hadron-hadron ones. It is a Leading Order (LO) generator and it calculates the Matrix Element (ME) at tree level, without virtual loop corrections.

It contains the Next-to-Leading-Order (NLO) QED corrections, thus it takes into account the irradiation of a photon from the decay muons. The real QCD corrections up to third order are taken into account as well.

Two samples have been generated: a sample in which the Z boson has been produced without QCD radiation and a sample in which the Z boson has been produced in association with real QCD radiation up to third order.

In Fig. 4.2 the di-muon generated mass and the Z boson mass are shown. The two lineshapes are fitted to a Breit-Wigner times an exponential component to describe the Z/γ^* interference term. The di-muon

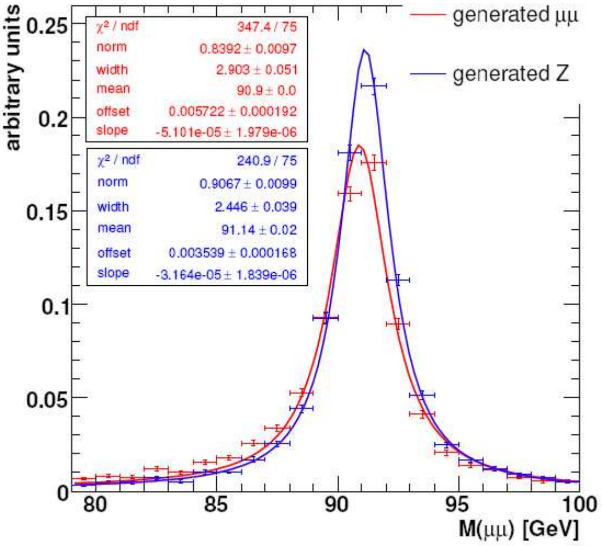


Figure 4.2: *Generated MC Z boson mass (blue) and di-muon mass (red). The result of the fit to two Breit-Wigner are reported. The difference between the two lineshapes is due to the Final State Radiation (FSR) and the convolution with the PDF.*

mass peak (in red) is about 50 MeV away from the nominal mass (in blue) because of the convolution with the PDF, while there is a further shift of about 250 MeV because of the QED final state radiation.

The samples have been then reconstructed in different conditions of the Silicon Strip Tracker: in ideal condition, i.e. the design one, and in a realistic detector scenario, which includes:

- an alignment scenario as foreseen after 10 pb^{-1} of collected data, which takes into account the alignment performance obtained during CRAFT data taking
- the mixing of the pure generated signal with the real tracker noise measured during CRAFT
- the exclusion of the tracker modules which will be inactive in the first data taking.

In the alignment scenario the detector modules in the simulation are displaced from their original position by a random amount within the uncertainty obtained after the alignment test performed during CRAFT data taking. The residual misalignment is of the same order of magnitude of the predicted uncertainty obtained after an alignment of the tracker performed with the collection of 10 pb^{-1} of data.

The mixing with the real noise is performed re-digitizing the Z signal switching off the noise in the simulation, and then mixing the "pure" signal with real noise events collected during CRAFT. The noise is obtained inverting a filter on the tracker pointing tracks skim, enlarging the exclusion region with respect to the tracker dimension. In this way the possible hits in the tracker are just noise hits, as no track at all is really pointing to it. The mixing is performed at the digi level after the Zero Suppression.

The exclusion of the tracker modules is based on the noise and bad channels identification described in Chapter 3. Some channels are known to be bad and not recoverable, so that they will be excluded from the data taking when LHC will start.

The result of the application of these three conditions in the simulation and reconstruction of the generated Z events gives an idea of how the first data might look like. For this reason Z events reconstructed in realistic conditions have been also used as a "fake data" sample to test the algorithm on a sample which mimics as closely as possible the first real data.

4.2.1 Muon reconstruction

The reconstruction of the Z boson is performed through the reconstruction of the two decay muons.

Muon reconstruction is performed in three stages:

- local reconstruction in muon spectrometer: the local pattern recognition starts from single hits to build track stubs (called segments) separately in each subdetector (CSC and DT);

- stand-alone reconstruction: the segments reconstructed in the muon chambers are used to generate "seeds" consisting of position and direction vectors and an estimate of the muon transverse momentum. These initial estimates are used as seeds for the track fits in the muon system, which are performed using segments and hits from DTs, CSCs and RPCs and are based on the Kalman filter technique. The result is an object referred to as "standalone muon" (STA). To improve the momentum resolution, a beam-spot constraint can be applied in the fit.

- global reconstruction: finally, standalone muon tracks are matched with tracker tracks to perform a global fit. For each standalone muon track, a search for a match with tracks among those reconstructed in the inner tracking system (referred to as "inner tracker tracks" - TRK) is performed, and the best-matching tracker track is selected. For each "tracker track" - "standalone muon" couple, the track fit is performed again with Kalman filter technique using all hits in both tracks. The result is an object referred to as "global muon" (GLB).

Tracker tracks and global muons have quite similar momentum and angle resolutions, since the tracker measurement totally dominates the precision on those quantities. The resolution on standalone muons the resolution is instead much worse. The pattern recognition in the muon spectrometer is, in fact, much more complex than in the tracker. There are many non-gaussian effects which can induce a scale bias and worsen the resolution: the multiple scattering and the energy loss in the material is big, there are regions where the magnetic field is strongly not uniform, finally, in case of showering, the wrong combination of hits can be chosen.

The Z bosons have been reconstructed in each event selecting the two opposite charge global muons in the CMS pseudorapidity range $|\eta| < 2.4$ whose invariant mass had the highest probability given by eq. 4.3 with respect to the expected mass and the resolution values estimated from the muon kinematics (Sec. 4.1). A cut on the minimum transverse momentum of the decay muons has also been established in 20 GeV/c because of the lack of statistics under this value, which caused problem in the likelihood minimization. The requirements just listed result in a selection efficiency of about 80%.

The muon momentum scale and the resolution measurement has been performed mainly using the measurement of the transverse momentum obtained from the inner tracker part of the the global muon. In order to cross check the results, the algorithm has been also run using the measurement of the momentum of the global track. This permits to check for a relative misalignment of muon chambers with respect to the inner barrel, or an incorrect modeling of the magnetic field in the outer toroids. The transverse momentum measured with the stand-alone muons has not been considered in the study because it is well known that their resolution is much worse.

4.3 Results

In this section the results of the application of the described algorithm are shown for the two different samples used. I will focus on the results obtained with the inner tracker tracks, leaving in Appendix the results regarding global muons. I will first show the results on the momentum scale calibration and the resolution measurement with the Z samples with and without QCD radiation reconstructed in ideal detector conditions, and successively with the samples reconstructed in the realistic scenario.

4.3.1 Ideal detector scenario

The first measurement has been done on the Z samples reconstructed in the ideal detector scenario. In this case one expects that the small biases which may act on the reconstructed muon momentum are just due to the difference between the parametrized Z boson probability density function generated with `HORACE` and the reconstructed quantities, which are subject to material interactions in the detector.

Before the likelihood fit can run on the set of data, it is necessary to choose an ansatz function to correct muon momenta. To help such a choice in the absence of any a priori knowledge of the possible sources of biases that may affect the data, the average mass can be studied as a function of the most basic kinematical variables: transverse momenta, rapidity and azimuth angles. The study is a very basic check of the existence of macroscopic effects.

In the case at hand, such a check does allow to spot very easily a slight rapidity dependence of the muon momentum scale, which is indeed attributable to differences in the amount of material crossed by the particles. As an ansatz thus, a function linear in transverse momentum and parabolic in the pseudorapidity has been chosen:

$$p'_T = [a_0 + a_1 \times p_T + a_2 \times |\eta| + a_3 \times \eta^2] \times p_T \tag{4.10}$$

The biases found on the muon scale for the two kinds of reconstructed muons (TRK, GLB) are listed in Table 4.1 for the sample with no QCD radiation and the one with real QCD radiation. The parameters values found using global muons and muon tracks in the fit are similar, accordingly with the fact that the tracker resolution is dominant for muons at the $M(Z)/2$ scale.

Par	TRK	GLB
a_0	0.9999±0.0017	0.998574±0.00000414
a_1	0.0	0.0
a_2	0.0	0.0
a_3	$(4.499\pm1.10)\times 10^{-4}$	$(6.499\pm0.023)\times10^{-4}$

(a) Ideal MC sample, no QCD correction

Par	TRK	GLB
a_0	1.00126±0.000084	1.0067±0.0027
a_1	$(-1.45586\pm0.0122)\times10^{-2}$	$(-2.694\pm0.404)\times10^{-4}$
a_2	$(1.188\pm0.099)\times10^{-3}$	0.0
a_3	$(3.34\pm0.258)\times10^{-1}$	0.0

(b) Ideal MC sample, QCD correction up to third order

Table 4.1: *Biases found on the momentum scale for TRK and GLB muons in the ideal detector scenario, using the ansatz functions of eq. 4.10. The values compatible with zero are indicated as zero.*

The effectiveness of the corrections obtained by the fit can be checked by studying the reconstructed mass as a function of the muon kinematics: the sample is divided in bins of the interesting kinematic variable and in each bin a lorentzian fit to the resonance peak is performed and the

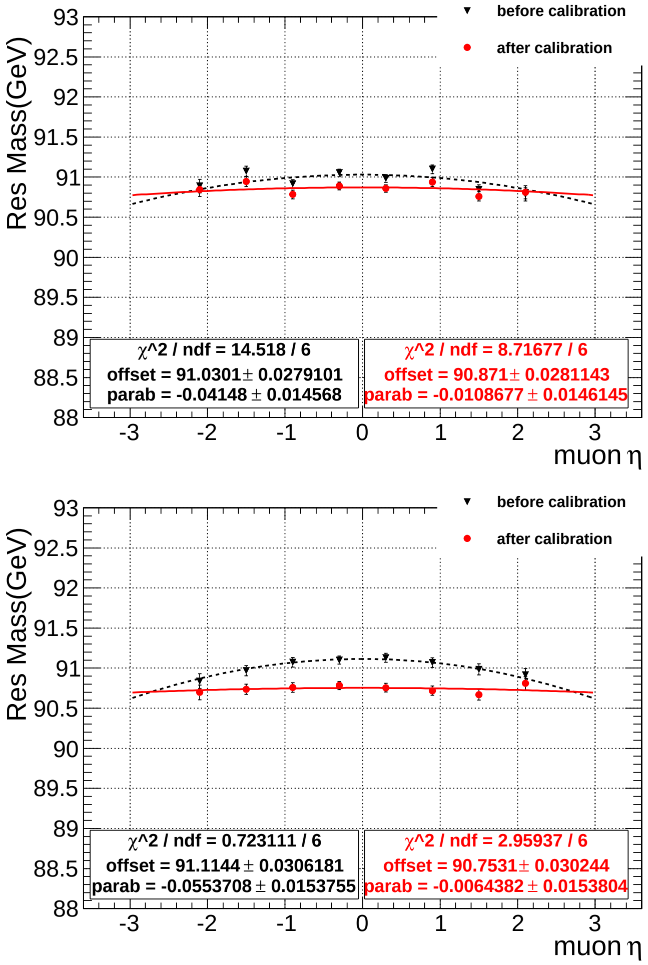


Figure 4.3: *Average invariant di-muon mass for Z without (top) and with (bottom) QCD radiation reconstructed in the ideal scenario with tracker muons as a function of the pseudorapidity, before and after the calibration of the muon momentum scale.*

mean of the fit is considered. The result of this study is presented in Fig. 4.3, where a profile of the average di-muon mass before and after the correction as a function of the rapidity is shown. The convention is to consider the bias corrected if the parabolic parameter (or whatever other bias parameter) after the momentum scale correction results to be compatible with zero within three standard deviations.

Here only the results obtained with TRK muons are shown. The same is shown for cross check in Appendix in Fig. 8 for global muons.

It should be noted that the parabolic dependence found by the fit of this plot is different from the bias computed with the calibration algorithm (reported in Table 4.1). This is because, in Fig. 4.3, an average on all the other muon kinematic variables is implicit, while the calibration algorithm computes the dependence of the scale on all the muon kinematic variables at once, taking into account cross-correlations. Thus this kind of plot, which will be widely used in the following to show the results, has to be interpreted just as a control plot to check that the algorithm is correcting biases in the scale. In the control fit of Fig. 4.3 it is possible to see that before the scale correction the parabolic term is clear and significative, while after the correction the parabolic term becomes compatible with 0. In this case the scale is considered to be corrected by the algorithm.

As in the calibration algorithm a complete probability density function is considered for the resonance, which also includes the Gaussian convolution as explained in Sec. 4.1, it is possible to fit the di-muon mass resolution together with the muon scale. This resolution is expressed, inside the likelihood, as a function of the resolution of the muon kinematic variables of eq. 4.8 so that the final results of the likelihood fit are values for the parameters which describe the resolution of these quantities. To this aim, meaningful ansatz functions and starting values of the parameters must be injected as input to the algorithm. A raw preliminary estimation of the muon resolution as a function of the muon kinematics is therefore needed.

In the ideal scenario, suitable forms for the $\sigma_i = G_i(\vec{x}; \vec{b}_j)$ of eq. 4.8 and initial values for their parameters $\vec{b}_j$ have been pre-determined with a Monte Carlo simulation of single muons originating from the CMS vertex with momenta ranging from 5 to 100 GeV/c, reasonably assuming that the resolution is not much different for single muons and for muons coming from the simulated resonance.

For the transverse momentum p_T the "relative" resolution is studied, i.e., the difference between reconstructed and generated p_T values divided by the generated one. For the azimuthal and polar angle the order of magnitude of the resolution is not expected to be directly dependent on the angle itself. Therefore the "absolute" resolution, i.e., the simple difference between reconstructed and generated values, is considered. The resolution is computed as the standard deviation of a Gaussian fit in each bin, in function of the muon kinematics. For the polar angle the resolution on $\cot\theta$ is considered. In fact, in the barrel region of the tracker, the longitudinal coordinate is measured by the silicon strips while the radial coordinate is fixed by the position of the sensitive modules. The measurement uncertainty is therefore proportional to the error on $\cot\theta$; the resolution on θ or η is instead a non-linear function of the measurement error.

The preliminary study performed on the simulation of single muons gave as result the following dependencies:

- p_T resolution linear as a function of p_T:

$$\sigma_{p_T} = (1.5 + 0.01 \times p_T) \times 10^{-2} \tag{4.11}$$

and parabolic as a function of η in different pseudorapidity regions. In detail two parabolas can be identified. One in the barrel region ($|\eta| \leq 1.1$) and one in the two endcap regions ($|\eta| > 1.1$). Nevertheless the parametrization has been done by point and not via a fit in the two separate regions in order to keep low the number of degrees of freedom in the likelihood fit. Two parabolas in fact have as a drawback a total of 6 parameters that need to be fitted to minimize the likelihood. The function parametrized by point has been demonstrated to have a good level of accuracy on the ideal scenario sample without increasing the number of degrees of freedom. The same function is not sufficient in the case of the realistic scenario, as it will be shown in next section.

- ϕ resolution parabolic as a function of η and with $1/\mathrm{p}_T$ dependence

$$\sigma_\phi = (1.1 + 0.22 \times \eta^2) \times 10^{-4} \tag{4.12}$$

$$\sigma_\phi = (0.065 + 3.6/p_T) \times 10^{-3} \tag{4.13}$$

- $\cot\theta$ resolution parabolic as a function of η and with $1/\mathrm{p}_T$ dependence

$$\sigma_{\cot\theta} = (4.0 + 0.78 \times \eta^2) \times 10^{-4} \tag{4.14}$$

$$\sigma_{\cot\theta} = (0.45 + 3.9/p_T) \times 10^{-3} \tag{4.15}$$

The results of the single muon simulation study are shown in Fig. 4.4. Each of the shown fits relies on an implicit average on all the other kinematic variables. Moreover the resolution of the different variables are considered separately although they are correlated: for instance the $1/p_T$ dependence of σ_ϕ and $\sigma_{\cot\theta}$ is already taken into account by σ_{pT}/p_T.

On the other side the muon calibration algorithm, described in Sec. 4.1, considers the resolution dependence on the full muon kinematic at once, not implying any average. It also fits the three resolutions ($\sigma_{\cot\theta}, \sigma_\phi, \sigma_{pT}/p_T$) together, through their impact on the di-muon mass resolution, therefore it takes into account the correlations between them avoiding any double counting.

After a "trial and error" procedure the following dependency was identified:

$$\sigma_{pT}/p_T = b_0 \times f(\eta, b_1) \tag{4.16}$$

where f(η, b_1) is the resolution function individuated with the single muon simulation study, which for each bin in η gives the resolution value as measured from the single muon simulation. The parameter b_1 is the value of the resolution in the last bin in η (corresponding to $|\eta| = 2.4$, which has been found to be variable and therefore it has been left as a free parameter for the likelihood minimization.

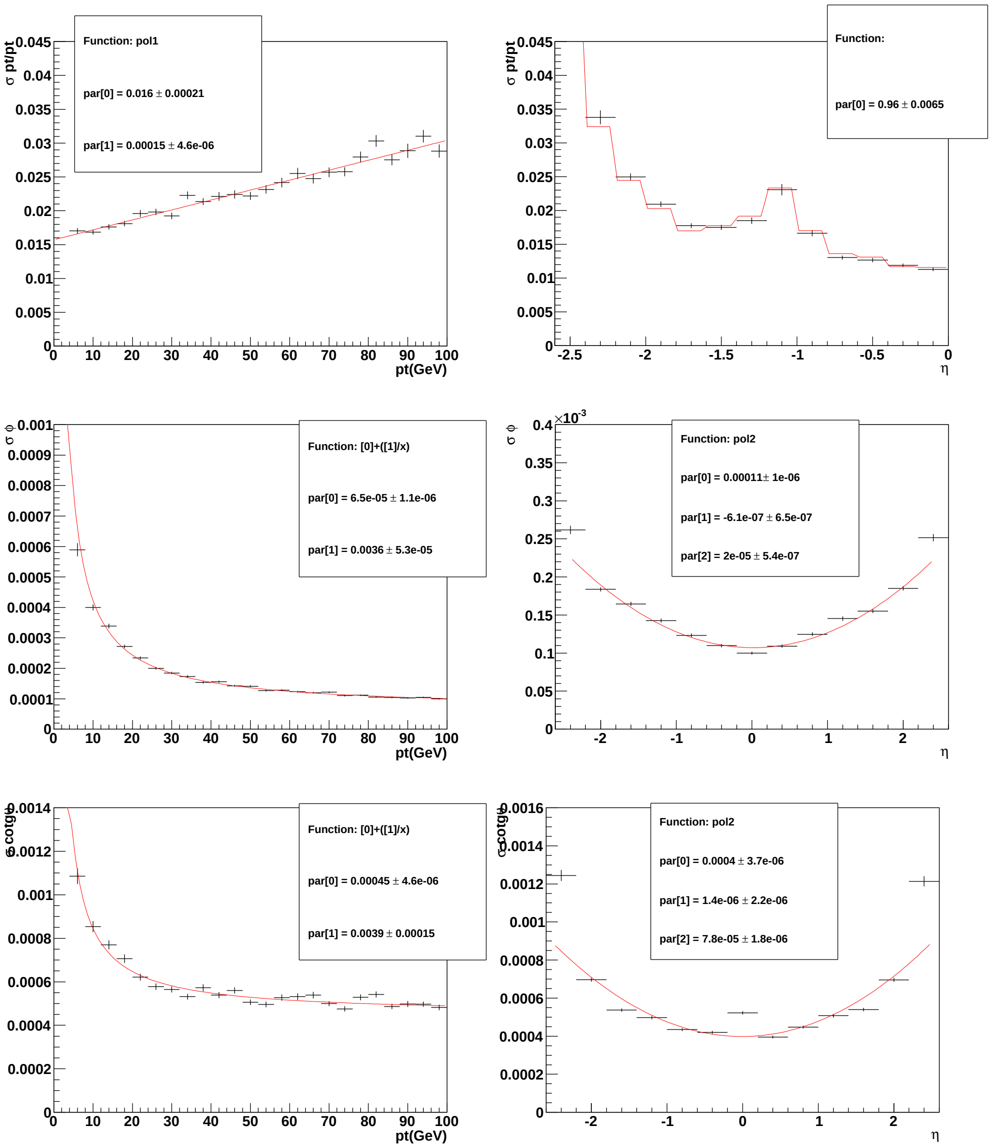


Figure 4.4: *Relative resolution on p_T (top), ϕ (middle), and $\cot\theta$ (bottom) as a function of muon p_T (left) and η (right) for global muons, as determined from Gaussian fits to the difference between true and reconstructed values in the single muon simulations (points with error bars). The overlaid hatched curves show the parametrization of the resolution used in the definition of the likelihood.*

The results of the fit to the resolution function are listed in Table 4.2 and shown in Fig. 4.5 for the Z sample without and with QCD radiation. The parameters of the resolution on the tracker tracks and the global muons are compatible within the errors. It should be noted also that b_0 is, in both cases, compatible with 1, which means that the resolution measured with muons from Z decay is the same as the one measured with the single muon simulation.

Par	TRK	GLB
b_0	0.952±0.033	0.953±0.052
b_1	0.0368±0.0058	0.049±0.0112

(a) Ideal MC sample, no QCD correction

Par	TRK	GLB
b_0	1.0721±0.0401	1.111±0.0513
b_1	0.0465±0.0088	0.0514±0.0095

(b) Ideal MC sample, QCD correction up to third order

Table 4.2: *Parameters of the resolution function for TRK and GLB muons in the ideal detector scenario, obtained fitting the ansatz function of eq. 4.16.*

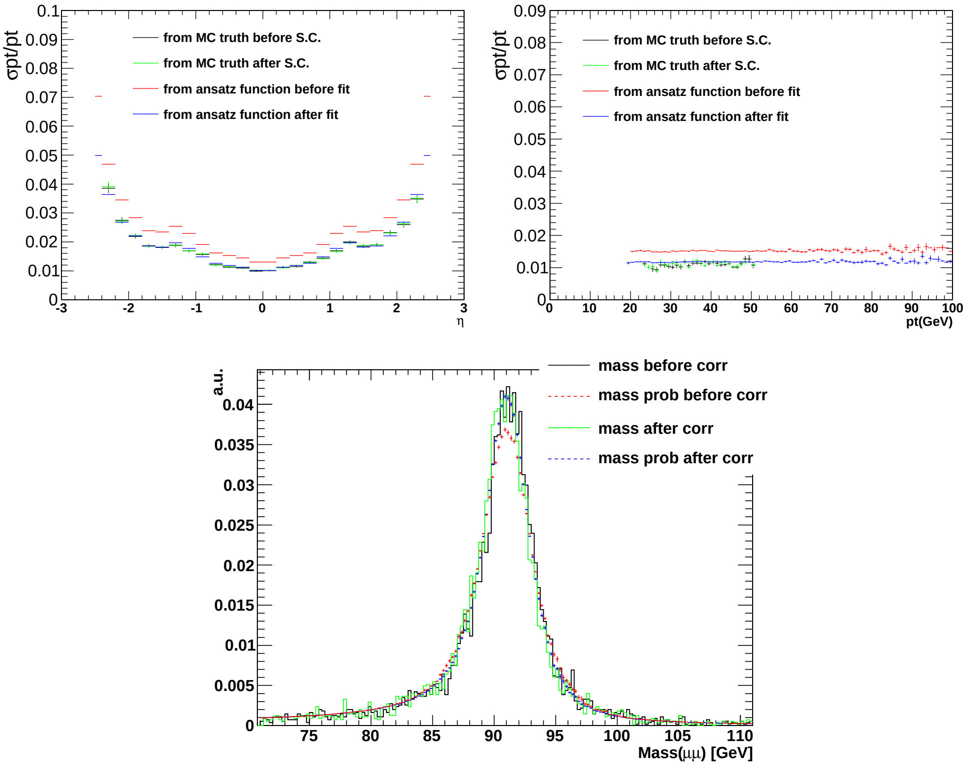


Figure 4.5: *Relative muon p_T resolution (σ_{pT}/p_T) as a function of the muon pseudorapidity (top left) and the muon transverse momentum (top right) for the Z sample with QCD radiation reconstructed in the ideal scenario, computed using inner tracker tracks. The comparison with relative resolution obtained from MC truth before and after the scale correction (S.C.) is also shown for cross check. The bottom plot shows the comparison of the Z mass line-shape and probability before and after the momentum scale correction and the resolution measurement.*

Fig. 4.5 shows the relative resolution on the transverse momentum as a function of the pseudorapidity and the transverse momentum itself, measured on the Z sample with QCD radiation using inner tracker tracks. Very similar results are obtained on the sample without QCD radiation and with global muons, shown in Appendix Fig. 12. The green curve represent the resolution obtained from the relative difference between the reconstructed p_T values and the generated ones; the red curve reproduces the ansatz function given to the algorithm before the fit and the blue curve is the result of the likelihood minimization. It is evident that the resolution function computed by the algorithm is in perfect agreement with the one computed using the reconstructed-generated comparison. The parabolic behavior as a function of the pseudorapidity is evident, with different shapes in the barrel and in the endcaps. In particular in the barrel the resolution has a value which goes from 0.01 to 0.02, while in the endcaps the resolution goes from 0.02 up to 0.04 in the very forward part of the detector. The behavior versus the transverse momentum is instead constant in the whole range.

The comparison between the reconstructed mass peak before (in black) and after (in green) the muon scale calibration is also shown. The red curve shows the probability computed with the initial values of the resolution parameters. In this case the resolution is overestimated and the red lineshape is wider. The blue histogram describes the probability after the resolution measurement, computed using the final fitted values of the parameters. After the fit the match with the Z lineshape measured is evident.

4.3.2 Realistic detector scenario

The same measurements presented in the previous section have been repeated on the Z samples reconstructed in realistic tracker conditions. The preliminary studies performed on the invariant di-muon mass as a function of the main kinematics variables show that the reconstruction in the realistic tracker scenario has led to a bias in the dependence of the mass versus the pseudorapidity and the azimuthal angle of the muon. In particular a parabolic bias as a function of pseudorapidity is expected also in this case, but it is found to be slightly stronger than in the ideal case. In addition a sinusoidal bias as a function of the azimuthal angle is found. As an ansatz thus, the following function has been chosen:

$$p'_T = (a_0 + a_4|\eta| + a_5\eta^2)p_T + \begin{cases} (a_1 sin(a_2\phi + a_3))p_T & q > 0 \\ (a_6 sin(a_7\phi + a_8))p_T & q < 0 \end{cases} \tag{4.17}$$

The biases found on the muon momentum scale for the TRK and GLB muons are listed in Table 4.3 for the two Z samples under study.

Par	TRK	GLB
a_0	$0.998205\pm2.15\times10^{-8}$	$0.964\pm2.2\times10^{-9}$
a_1	$0.00248293\pm1.177\times10^{-7}$	$(1.104\pm0.088)\times10^{-3}$
a_2	$1.05863\pm6.96\times10^{-5}$	1.
a_3	0.0	2.970 ± 0.014
a_4	$(-1.74854\pm0.000642)\times10^{-4}$	$(-6.414\pm0.017)\times10^{-3}$
a_5	$(-1.0134\pm0.000041)\times10^{-4}$	$(2.61964\pm0.010)\times10^{-3}$
a_6	$(5.67\pm0.00076)\times10^{-4}$	$(7.75957\pm0.000018)\times10^{-5}$
a_7	$1.0167\pm5.61\times10^{-6}$	0.154982 ± 0.0012
a_8	0.0	1.7134 ± 0.001199

(a) Fake data sample, no QCD correction

Par	TRK	GLB
a_0	0.954 ± 0.0042	0.996 ± 0.0015
a_1	0.0	0.0042 ± 0.0013
a_2	-0.961 ± 0.088	1.240 ± 0.142
a_3	2.0021 ± 0.0501	0.0
a_4	0.0144 ± 0.00465	0.0021 ± 0.00035
a_5	-0.00498 ± 0.000452	$-(5.54\pm0.16)\times10^{-4}$
a_6	0.0936 ± 0.00183	$-(6.28\pm0.15)\times10^{-4}$
a_7	-0.3202 ± 0.0652	1.271 ± 0.205
a_8	1.589 ± 0.082	0.0

(b) Fake data sample, QCD correction up to third order

Table 4.3: *Biases found on the momentum scale for TRK and GLB muons in the realistic detector scenario, using the ansatz function of eq. 4.17. The values compatible with zero are indicated as zero or not reported at all.*

In Fig. 4.6 the profile of the invariant di-muon mass as a function of the pseudorapidity and of azimuthal angle is shown for the Z sample reconstructed in realistic condition, as control plots to check that the scale has been corrected. In this case the parabolic bias versus η is slightly stronger (more than double with respect to the ideal case), nevertheless it is removed by the calibration procedure, as it can be verified looking at the parabolic parameters of the fit before and after the correction. The bias as a function of the azimuthal angle is also evident and it is corrected after the application of the algorithm.

The preliminary study on the resolution in order to search for a suitable ansatz function and for reasonable starting values has been repeated for the realistic scenario in a similar way with respect to the study done in the ideal scenario, considering the resolution obtained from the reconstructed and generated kinematical variables of the two decay muons.

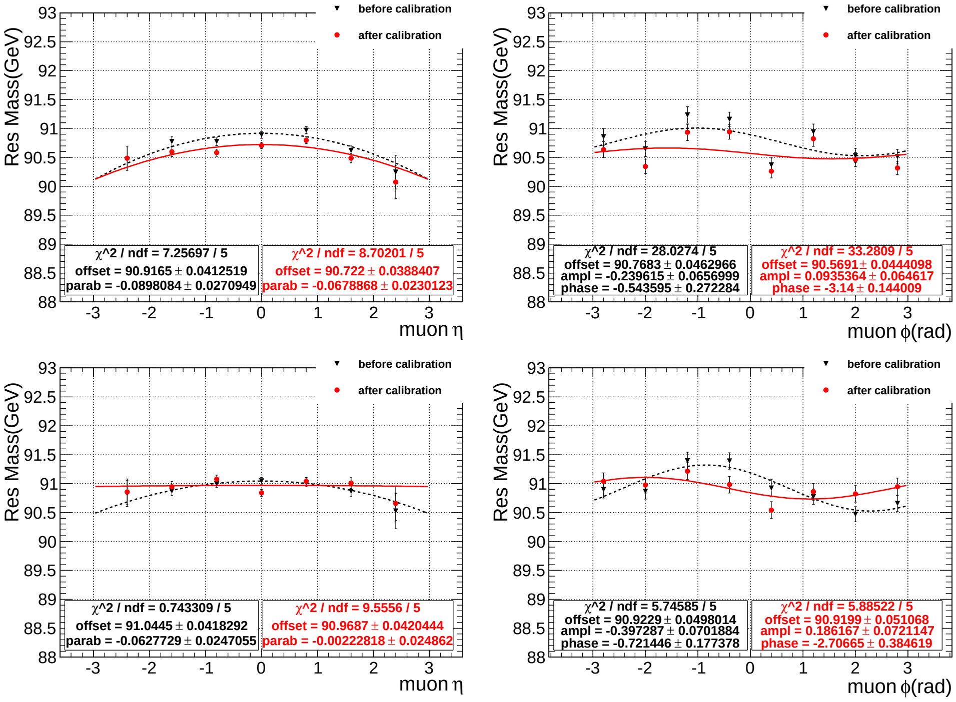


Figure 4.6: *Invariant di-muon mass of Z without (top) and with (bottom) QCD radiation reconstructed in the realistic detector scenario, computed with inner tracker tracks as a function of the pseudorapidity (left column) and azimuthal angle (right column), before and after the calibration of the muon momentum scale. The same is shown in Appendix Fig. 9 for global muons.*

After this preliminary study quite a complex resolution function has been identified, defined in a two-dimensional (p_T, η) space:

$$\sigma_{pT}/p_T = b_0 \times p_T +$$
$$+ \begin{cases} b_1 \times \eta^2 & |\eta| > 0.6 \ \& \ |\eta| < 1.3 \\ b_2 \times (|\eta| - b_3)^2 & \eta > 1.3 \\ b_4 \times (|\eta| - b_5) + b_6 \times (|\eta| - b_5)^2 & \eta < \text{-}1.3 \end{cases} \tag{4.18}$$

Plugging this expression into the algorithm as ansatz function with reasonable starting values, the results shown in Table 4.4 are obtained for the parameters b_i for tracker tracks and global muons. Also in the realistic scenario the resolution measured using the estimation of the transverse momentum from global muons or inner tracker tracks give comparable results. In most of the cases they are compatible within three standard deviations.

Par	TRK	GLB
b_0	$(4.164\pm0.37)\times10^{-4}$	$(4.0211\pm0.0129)\times10^{-4}$
b_1	0.0	$(5.121\pm0.0515)\times10^{-3}$
b_2	0.116 ± 0.016	0.1015 ± 0.00036
b_3	1.364 ± 0.0479	1.35174 ± 0.00121
b_4	0.0	0.0
b_5	1.360 ± 0.078	1.386 ± 0.00158
b_6	0.071 ± 0.0145	0.0775 ± 0.0004

(a) Fake data sample, no QCD correction

Par	TRK	GLB
b_0	$(4.013\pm0.296)\times10^{-4}$	$(3.76\pm0.3)\times10^{-4}$
b_1	$(5.67568\pm1.606)\times10^{-3}$	0.0
b_2	0.0688 ± 0.0098	0.070 ± 0.014
b_3	1.114 ± 0.063	1.2045 ± 0.131
b_4	0.0	0.0
b_5	1.455 ± 0.052	1.444 ± 0.076
b_6	0.103 ± 0.018	0.087 ± 0.016

(b) Fake data sample, QCD correction up to third order

Table 4.4: *Parameters of the resolution function on transverse momentum for TRK and GLB muons obtained in the realistic detector scenario, using the ansatz function of eq. 4.18. The values compatible with zero are indicated as zero or not reported at all.*

The relative resolution on transverse momentum obtained from the preliminary study, the ansatz function before the fit and the final resolution function resulting from the fit are shown in Figures 4.7 for tracker tracks as a function of the pseudorapidity and the transverse momentum.

The resolution on the transverse momentum in the realistic scenario is on average higher than in the ideal scenario, as expected. In particular in the barrel its value is in the interval 0.02-0.03, while in the endcaps it increases faster than in the ideal case, going from a minimum value of 0.03 to even values of 0.09.

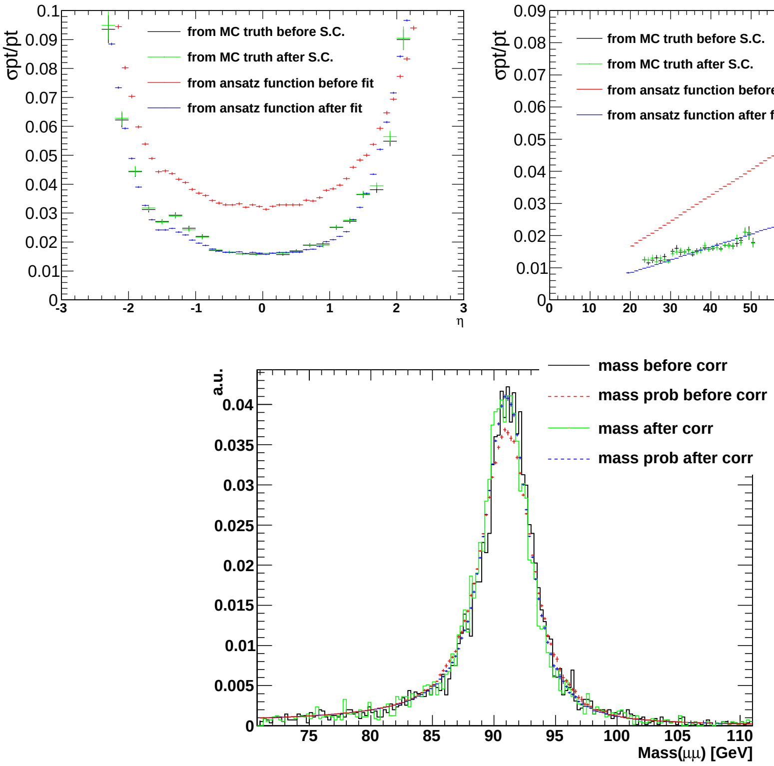


Figure 4.7: *Relative muon p_T resolution (σ_{pT}/p_T) as a function of the muon pseudorapidity (top left) and the muon transverse momentum (top right) measured on the Z sample with QCD radiation reconstructed in the realistic scenario, computed using inner tracker tracks before and after the scale correction. The comparison with relative resolution obtained from MC truth before and after the scale correction (S.C.) is also shown for cross check. The bottom plot shows the comparison of the Z mass lineshape and the probability before and after the momentum scale correction and the resolution measurement.*

The comparison between the reconstructed mass peak before (black) and after (green) the muon scale calibration is also shown. Also in this case it is evident the good match of the probability after the fit (in blue) with the measured Z lineshape, with respect to the probability before the fit (in red) which represents a wider Z peak due to the overestimated resolution.

4.4 Monte Carlo simulation resolution tuning

As shown in the previous section, already with few pb^{-1} of integrated luminosity a measurement of the resolution on the transverse momentum will be feasible directly from data, from the Z boson line shape. At the same time, it has been shown how different the resolution can be with a sample reconstructed simulating a detector in ideal condition from the resolution measured on the same sample reconstructed simulating realistic conditions (i.e. possible residual misalignments and miscalibrations). On the other hand it is very important for precision measurement and for early discoveries as well, that the Monte Carlo simulation is able to reproduce the data with good approximation. A method is therefore necessary to tune the Monte Carlo simulation of the CMS detector with early data (10 pb^{-1}) in order to have the most realistic description of the detector and to obtain from the Monte Carlo the most precise description of data.

In this section a method is presented to correct the Monte Carlo resolution on transverse momentum in order to achieve a good agreement with the resolution measured from early data. For the development and test of the method the two samples already introduced have been used. The Z samples reconstructed in the ideal scenario represent the ideal Monte Carlo sample, on which one wants to apply the method in order to achieve the same resolution measured on data. The Z samples reconstructed in the realistic detector scenario constitute a "fake data" sample, which mimics a real data sample after the collection of 10 pb^{-1} integrated luminosity.

4.4.1 Correction method

The main hypotheses on which the method here presented relies are the following:

- all the effects introduced in the track reconstruction by a detector in realistic condition with respect to a detector in ideal condition can be summarized simply in an additional Gaussian smearing of the transverse momentum;
- the Gaussian function which smears the generated (or real) transverse momenta during the reconstruction in an ideal detector and the additional Gaussian smearing of the reconstructed transverse momenta because of miscalibrations and misalignment can be treated as two uncorrelated Gaussian effects.

Under these two hypotheses one can state that the resolution on real data can be obtained just as the quadratic sum of the resolution measured in the ideal simulation and the additional smearing due to misalignment and miscalibrations:

$$\sigma^2_{data} = \sigma^2_{MC} + \sigma^2_{add} \tag{4.19}$$

where σ_{data} and σ_{MC} are the p_T resolutions measured on the "fake data" and on the ideal Monte Carlo sample respectively. Eq. 4.19 allows to parametrize the additional smearing as a function of muon kinematics (transverse momentum and pseudorapidity in particular, or whatever variable on which the resolution is found to depend on). Assuming that the additional term introduces random biases on the kinematic measurement, the smearing on the muon momentum of the ideal sample can be done using a Gaussian function. The transverse momentum changes as:

$$p_{Tsmear} = p_{TMC} \times Gaus(1, \sigma_{add}) = p_{TMC} \times Gaus(1, \sqrt{\sigma^2_{data} - \sigma^2_{MC}}) \tag{4.20}$$

This kind of smearing takes into account all the effects introduced at once, as they are convoluted and not separable one from the other in the real data. The aim is to parameterize the residual differences that can show up between the Monte Carlo and the data once the main effects introduced by known problems and defects (i.e. residual miscalibration and misalignment) have been already addressed.

4.4.2 Results

For this study only the results obtained on the Z sample with QCD correction and on inner tracker tracks will be shown, as it has been seen that no appreciable differences are present in the different samples and between the momenta measured on the global or the inner tracks.

Once the ideal Monte Carlo sample has been modified smearing the transverse momenta with eq. 4.20, a study of the resolution as relative difference between the reconstructed and the generated values of the transverse momentum has been done. It can be seen from Fig. 4.8 that there is a perfect agreement between the resolution measured on the smeared Monte Carlo sample and that on the "fake data" sample, both in function of the pseudorapidity and of the transverse momentum. This is a first check that the method is working properly.

Nevertheless, a more accurate estimate of the resolution can be obtained running the algorithm for the momentum scale correction and the resolution measurement already used in the previous section.

For the scale correction the same function used for the ideal Monte Carlo sample has been used, as no bias is expected from a gaussian smearing with unitary mean. The results for the scale correction (Table 4.5 and Fig. 4.9 with the usual control plot) indeed show that no bias is evident in the scale already before the scale correction, both as a function

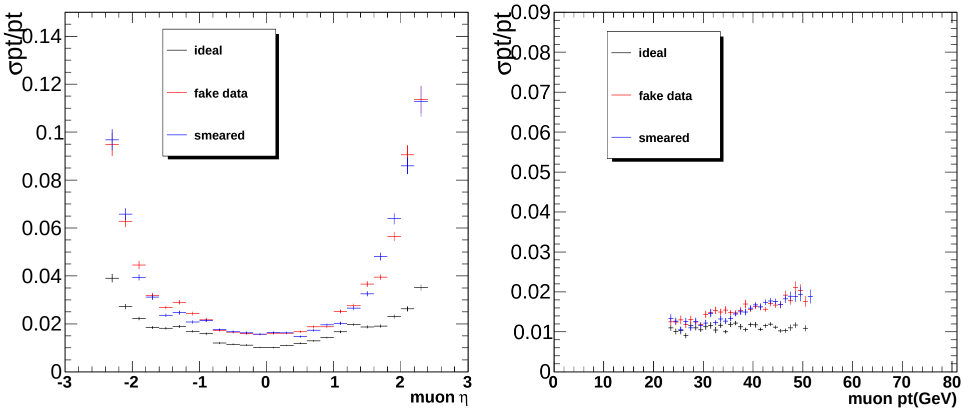


Figure 4.8: *Relative muon p_T resolution (σ_{pT}/p_T) as a function of the muon pseudorapidity (left) and the muon transverse momentum (right) for the three Z data samples: ideal MC, fake data and smeared MC, computed as relative difference between reconstructed transverse momentum and generated value, using inner tracker tracks.*

of the pseudorapidity and as a function of the azimuthal angle. The first parameter is compatible with 1 and the remaining three parameters are compatible with 0. This is expected for two main reasons. First of all, as already stated, the gaussian smearing is not expected to apply any systematic bias to the momenta; in addition, the values of the parameters used for the smearing function had been obtained correcting the momentum scale. Therefore the smearing should already take into account the momentum scale correction.

Par	TRK
a_0	1.00224±0.0025
a_1	(-6.089±4.6)×10^{-4}
a_2	(1.966±3.476)×10^{-3}
a_3	(-1.271±1.660)×10^{-3}

Table 4.5: *Biases on transverse momentum for TRK muons computed on the smeared Monte Carlo sample, using the same ansatz functions used for ideal data (eq. 4.10). Most of the values are compatible with zero or one, as expected from the fact the momenta are smeared using a gaussian function with mean = 1 and that the scale had already been corrected before the smearing.*

Simultaneously to the scale correction parameters, the resolution parameters are calculated by the likelihood minimization. The results are reported in Table 4.6, together with the parameters obtained in the previous section for the ideal Monte Carlo sample (ideal detector condition) and for the “fake data” sample (realistic detector condition). It can be

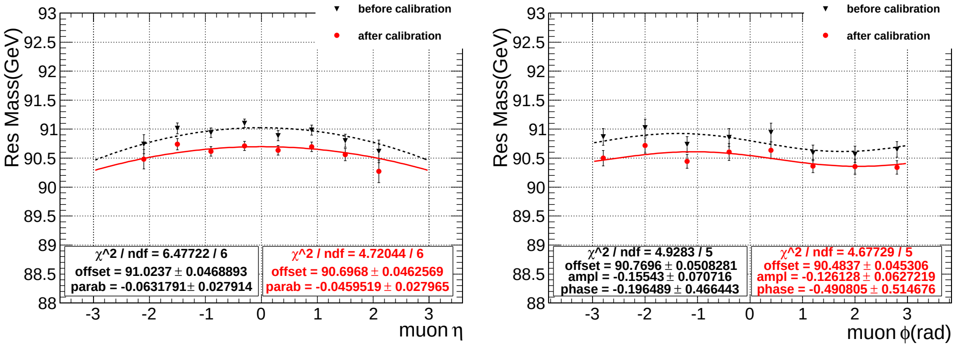


Figure 4.9: *Invariant di-muon mass for Z with QCD radiation computed on the smeared Monte Carlo sample with inner tracker tracks as a function of the pseudorapidity (top) and of the azimuthal angle (bottom), before and after the calibration of the muon momentum scale.*

seen that the parameters found from the smeared sample and the ones obtained on the "fake data" sample are compatible with each other within three standard deviations. In particular the b_0 parameter, which represents the dependence of the resolution on the transverse momentum, shows an excellent agreement between the "fake data" and the smeared Monte Carlo sample. Only two parameters (b_2 and b_3) show a different behavior in the smeared sample with respect to the "fake data" sample. These correspond to the resolution parameters in the endcap region with $|\eta| > 1.3$, in which the parabolic shape is difficult to estimat. Nevertheless the shape of the p$_T$ resolution computed by the algorithm as a function of η and p$_T$ well reproduces the resolution measured from the comparison of reconstructed and generated transverse momentum, as shown in Fig. 4.10. Here the relative resolution on transverse momentum as a function of pseudorapidity and transverse momentum is shown before and after the likelihood minimization, together with the relative resolution computed from the Monte Carlo truth for cross check.

Parameters	Values		
	Ideal	Fake	Smeared
b_0	1.0721±0.0401	$(4.013\pm0.296)\times10^{-4}$	$(4.020\pm0.535)\times10^{-4}$
b_1	0.0465±0.0088	$(5.67568\pm1.606)\times10^{-3}$	0.0
b_2		0.0688±0.0098	0.0077±0.00405
b_3		1.114±0.063	-3.827±0.347
b_4		0.0	0.0
b_5		1.455±0.052	1.337±0.120
b_6		0.103±0.018	0.0726±0.018

Table 4.6: *Parameters of the resolution function on transverse momentum for TRK muons computed on the smeared Monte Carlo sample, using the same ansatz functions used for fake data (eq. 4.18). The values compatible with zero are indicated as zero or not reported at all.*

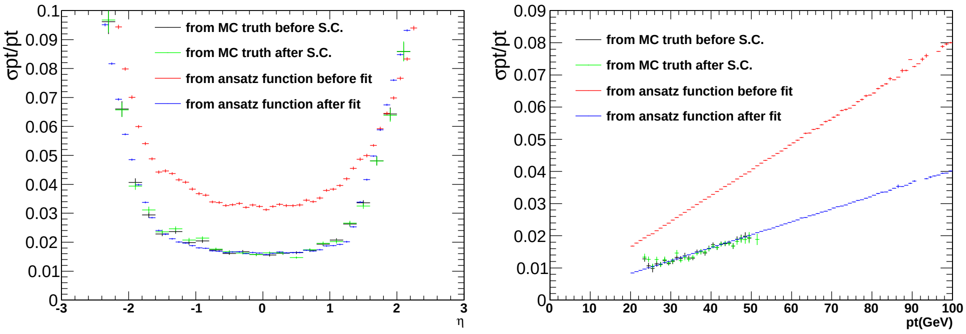


Figure 4.10: *Relative muon p_T resolution (σ_{pT}/p_T) as a function of the muon pseudorapidity (left) and the muon transverse momentum (right) for the smeared Monte carlo sample of Z with QCD radiation, computed using inner tracker tracks. The comparison with relative resolution obtained from MC truth is shown.*

In the smeared sample the resolution on the transverse momentum results to be of the same order of magnitude with respect to the "fake data" sample: ranging from 0.018 to 0.03 in the barrel region ($|\eta| \leq 1.4$), and from 0.028 to 0.09 in the endcaps.

The comparison between the resolution on the transverse momentum measured with the likelihood minimization from the Monte Carlo sample, the fake data sample and the smeared MC sample is shown in Fig. 4.11, as a function of the muon kinematics (pseudorapidity and transverse momentum).

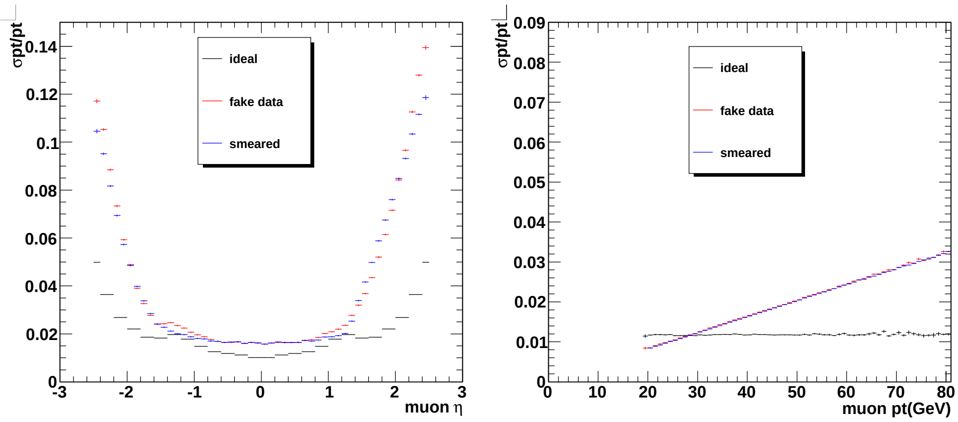


Figure 4.11: *Relative muon p_T resolution (σ_{pT}/p_T) as a function of the muon pseudorapidity (left) and the muon transverse momentum (right) for the three Z data samples: ideal MC, fake data and smeared MC, computed using inner tracker tracks.*

After the smearing the resolution of the Monte Carlo is in perfect agreement with the resolution on the "fake data" sample. The agreement is clear also if one looks at the Z lineshape for the three samples, shown in Fig. 4.12. The three lineshapes are fitted to a Breit-Wigner times an exponential function to take into account the Z/γ^* interference. The Z lineshape obtained from the "fake data" sample and the one obtained from the smeared Monte Carlo sample show a larger width, with respect to the lineshape measured with the ideal Monte Carlo, which is due to the worse resolution. The relative differences among the three lineshapes are also shown. The effect of the smearing is instead less evident in the Z transverse momentum distribution, where the differences between the realistic detector scenario and the ideal one does not affect the measurement. This is another good result, as it shows that the smearing of the muon transverse momentum does not modify Z kinematical variables which are actually not affected by the reconstruction in the realistic scenario.

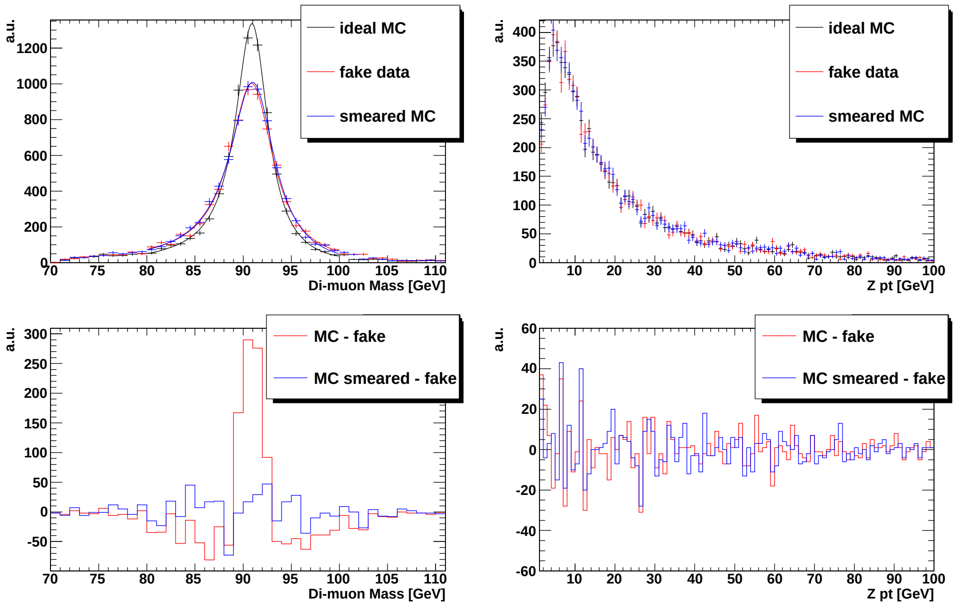


Figure 4.12: *Z invariant mass and transverse momentum distributions for the three data samples: ideal MC, fake data and smeared MC (top) and relative differences (bottom).*

The results obtained measuring the resolution on the smeared sample are in excellent agreement with the one obtained on the "fake data" sample. This is a considerable result, which will enable to tune the detector simulation in the Monte Carlo in the early phase of the data taking as soon as a statistically relevant sample of $Z \rightarrow \mu\mu$ events will be collected. The results achieved on the Z boson are applicable to any other physics analysis, hence constituting a milestone for the Standard Model physics at CMS, and opening the route to new physics.

Summary

This thesis describes the tracking system in all its aspects: from the calibration and the local reconstruction tested during cosmic Global Runs, to the effect that possible residual miscalibrations and misalignments can have on the physics measurements with first real data.

The Silicon Strip Tracker, with more than 15 000 silicon modules, is the most important part of the tracking system in CMS, together with the Pixel detector. A careful calibration procedure is needed in order to have the best estimate of the position of the particles passing through the detector and thus to measure the transverse momentum in the most accurate way. The measurement of particles transverse momentum is of fundamental importance for most of the physics analysis with the CMS detector at the LHC.

I took part into the calibration of this complex object, developing Data Quality Monitoring tools in order to check the performance of the detector in terms of all the relevant quantities.

During the last three years of data taking the Silicon Strip Tracker showed excellent performances, with less than 1% of dead channel, a very stable signal-to-noise ratio (well above 25) and an hit reconstruction greater than 99.5%. The track reconstruction efficiency has been measured with different methods and it has also resulted to be always above 99%.

The whole tracker went through alignment procedures based on different algorithms to validate and cross check the results: this allowed to achieve a precision on module position measurement of about 3 μm in the barrel and less than 8 μm in the endcaps, respectively. This alignment is compatible with the maximum statistical precision that can be achieved using cosmic tracks, which do not come from the interaction point. Thus residual systematic effects survive and will need collision data to be completely removed.

The cosmic data taking has constituted a milestone in the calibration of the CMS detector towards final commissioning with colliding beam data and it showed that design performances can be achieved.

The LHC is now entering collision mode, after almost 20 years of R & D and commissioning. Soon, a large amount of data will become available to perform the commissioning with collision data and to finally start the physics analysis. One of the first measurements that will become possible is the one of the resolution on the transverse momentum, which is a fundamental quantity to be measured with the best possible precision at CMS for all physics studies. I have illustrated an algorithm which permits to perform this measurement with the first 10 pb^{-1} integrated luminosity, taking advantage of the well known Z boson line-shape.

The algorithm has been used on Monte Carlo samples reconstructed in an ideal scenario in which the silicon tracker is perfectly aligned and calibrated, and in a realistic scenario which takes into account the residual miscalibrations and misalignments after cosmics data taking. This so-called “fake data” sample mimics a real data sample collected with the quoted integrated luminosity.

The resolution measurement performed with this algorithm has shown excellent results, perfectly reproducing the resolution obtained by comparing reconstructed and generated quantities. In the ideal scenario the relative resolution on transverse momentum has resulted to be about 0.015 in the barrel and slightly higher in the endcaps, showing a constant behavior as a function of p_T. In the realistic scenario the resolution measurement returns values ranging from 0.018 to 0.03 in the barrel and from 0.03 to 0.09 in the endcaps. This much higher value in the endcaps is in agreement with the lower precision achieved with the alignment in the forward regions, because of the lack of cosmic rays at such a high polar angle. Moreover a linear dependence on the transverse momentum has been found.

The measurement of the resolution on the two Z samples has been used to find a method to force the Monte Carlo simulation to show the same resolution as in real data. A gaussian smearing of the transverse momentum has been found to fulfill this requirement, incorporating all at once the possible residual inaccuracies in the simulation and the reconstruction which have not been addressed with the calibration procedures.

The smearing method is shown to successfully correct the resolution on the ideal Monte Carlo sample already with the first 10 pb^{-1} integrated luminosity. The resolution measured on the smeared Z sample shows an excellent agreement with the one obtained on the “fake data” sample. This is an outstanding result that will be achievable as soon as a sufficient sample of Z $\rightarrow \mu\mu$ events will be collected. The tuning of the Monte Carlo simulation in the very early phase of the data taking will allow a consistent comparison with real data, opening the route to new physics discoveries.

Appendix

In this Appendix some more results are shown regarding all the studies reported in this thesis.

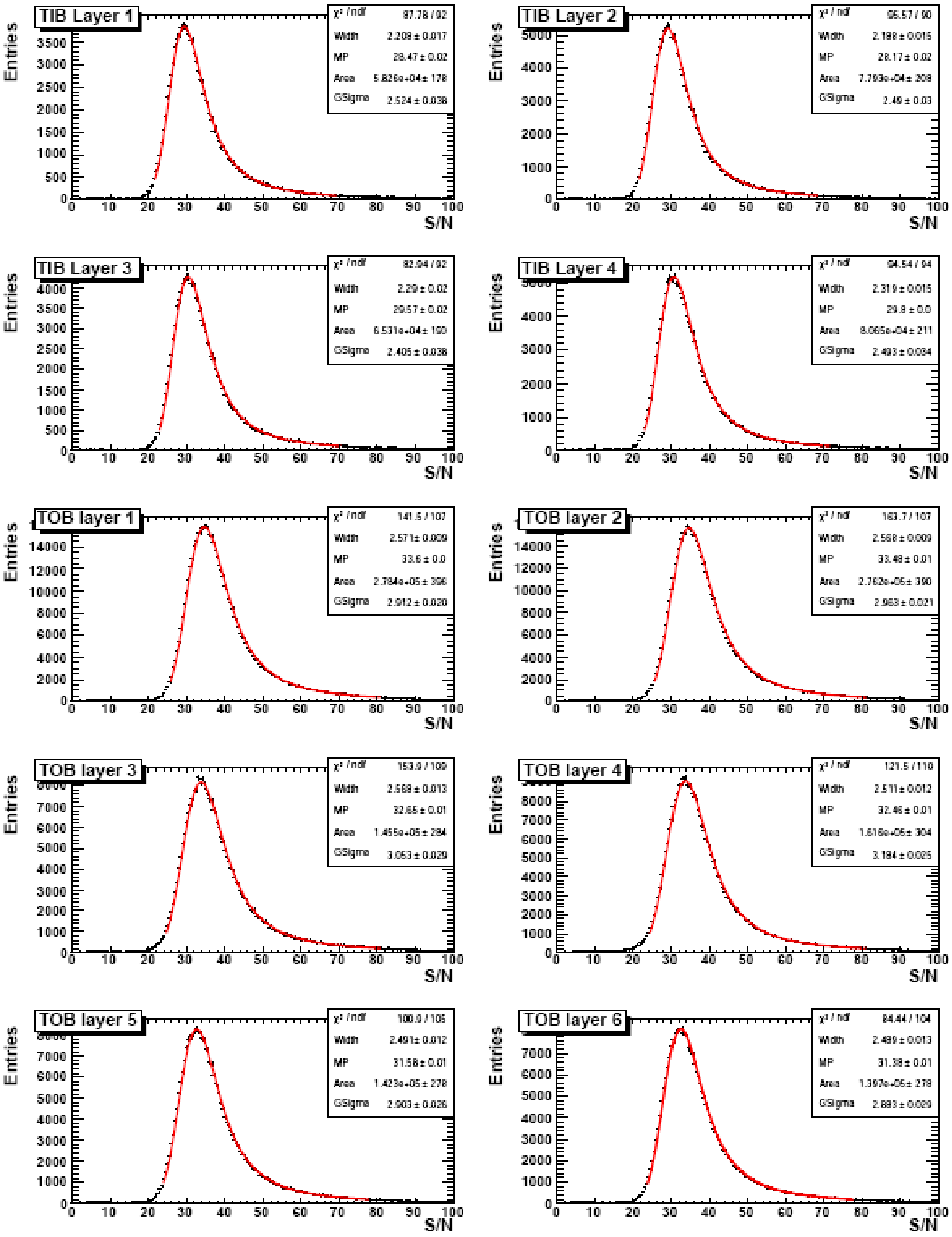


Figure 1: *The signal-to-noise corrected for the track angle for TIB and TOB layers at T = -10° C.*

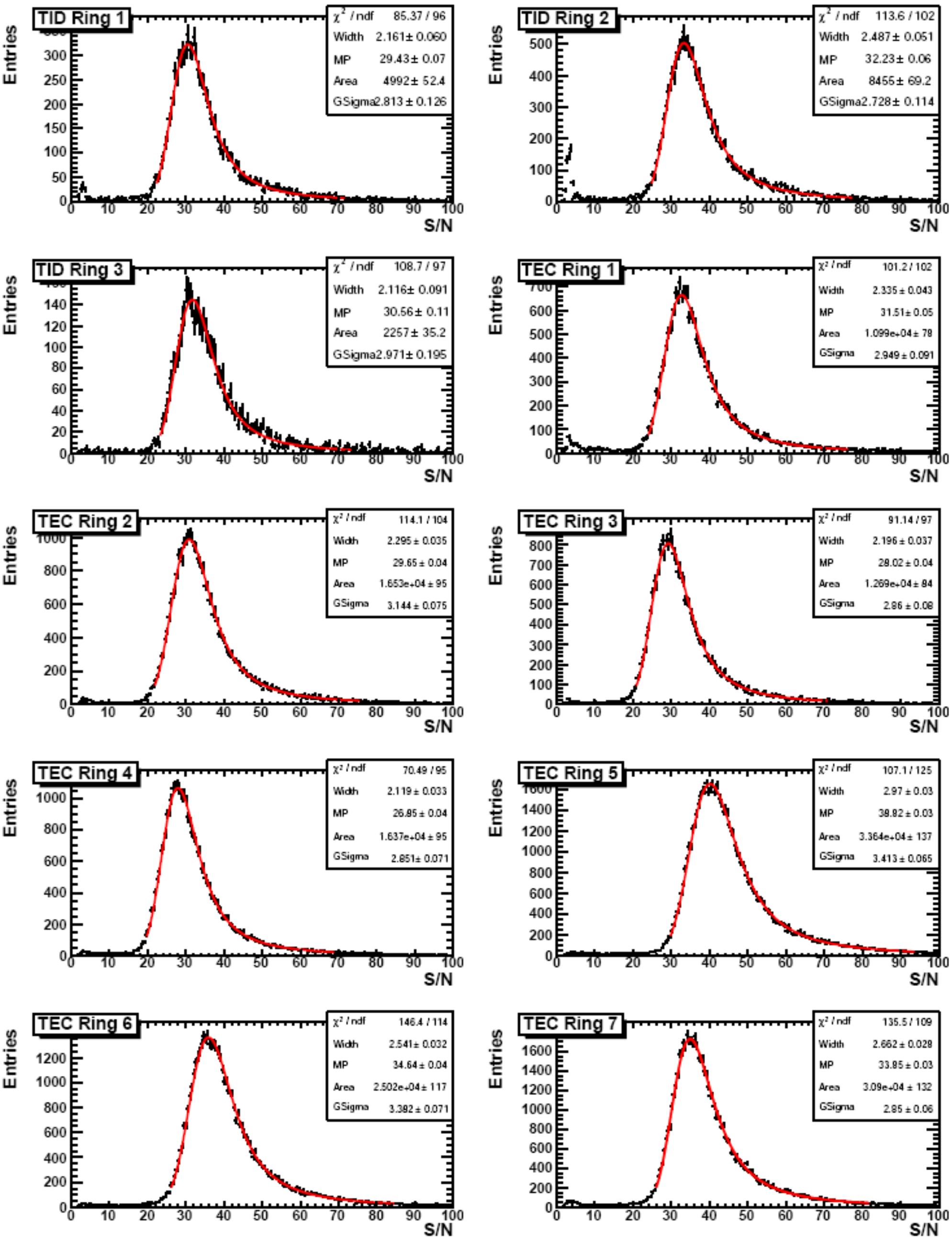


Figure 2: *The signal-to-noise corrected for the track angle for TID and TEC rings at T = -10° C.*

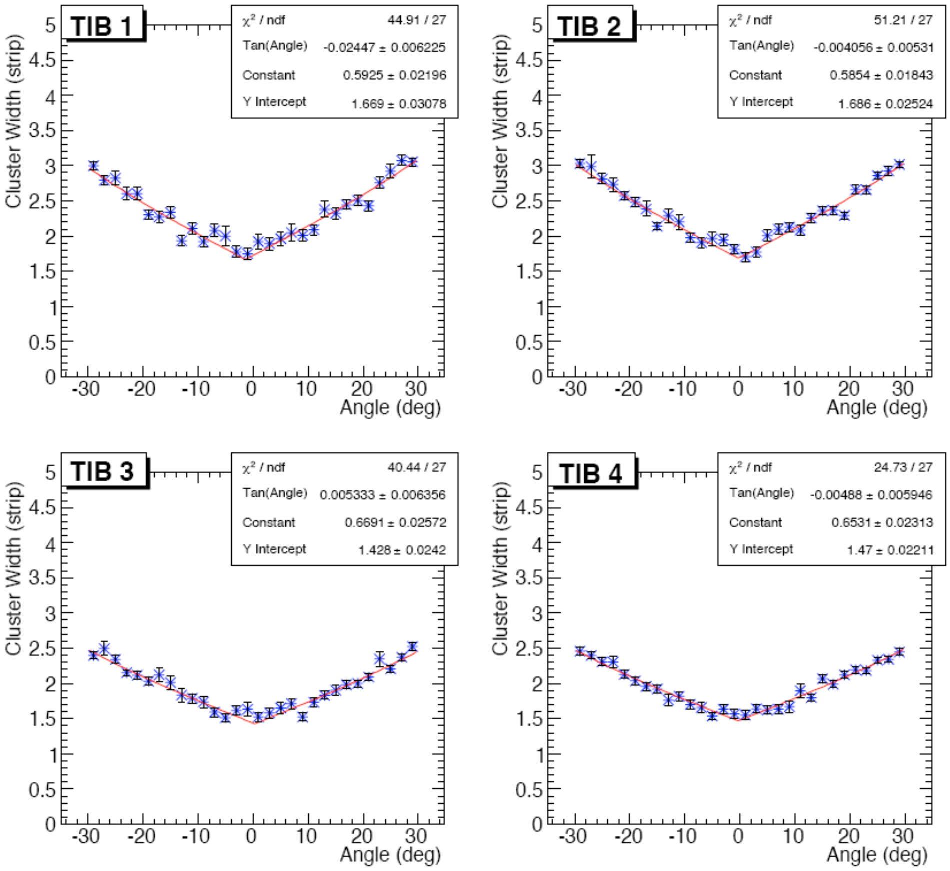


Figure 3: *Cluster width versus θ_t for TIB layers.*

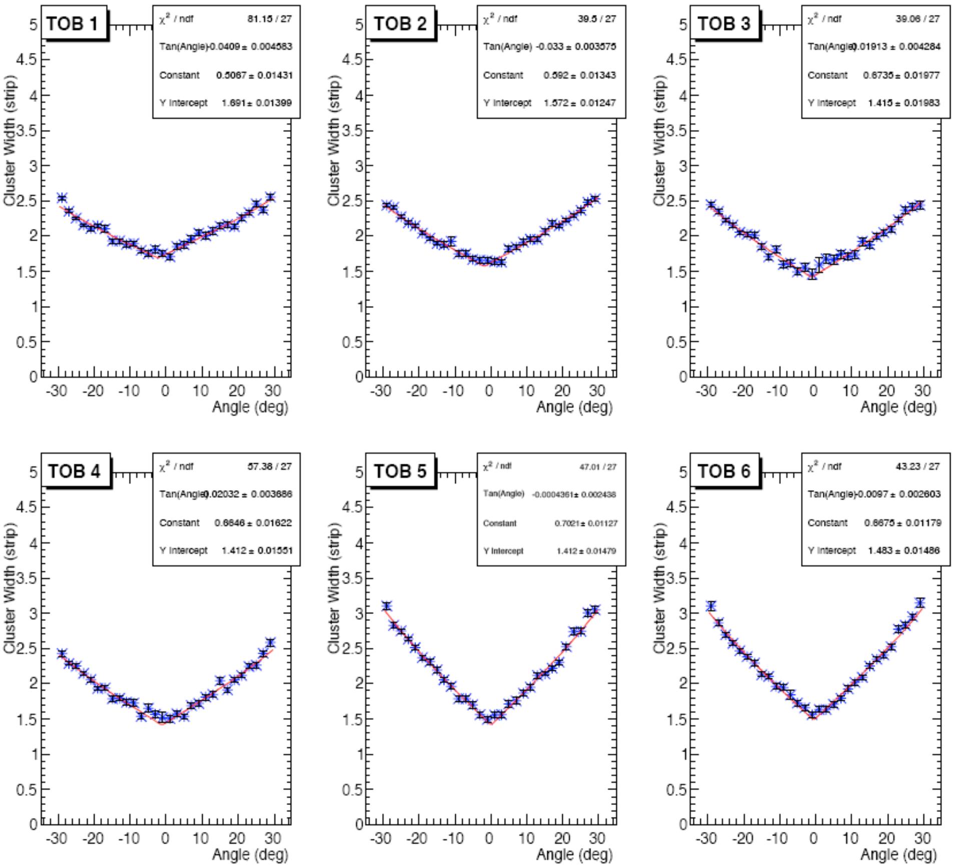


Figure 4: *Cluster width versus* θ_t *for TOB layers.*

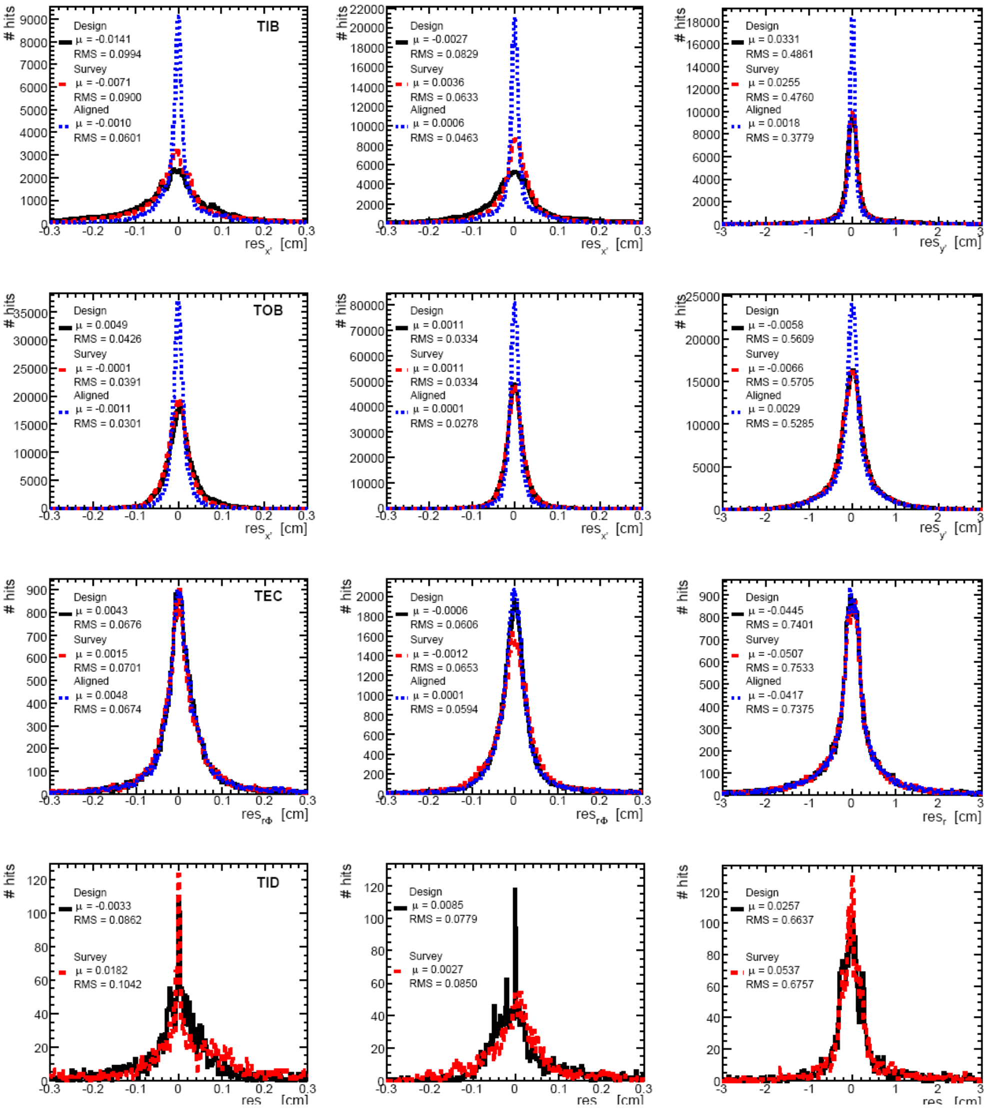


Figure 5: *Hit residuals for different geometries: ideal (solid/black), survey (dashed/red), and track-based alignment (dotted/blue, HIP). Four Tracker sub-detectors are shown in the top row (TIB), second row (TOB), third row (TEC), and bottom row (TID). The absolute local x'-residuals are shown for single-sided modules (left) and double-sided modules (middle), while local y'-residuals are shown for the double-sided modules only (right). For the endcap modules (in TEC and TID) transformation to the $r\phi$ and r residuals is made.*

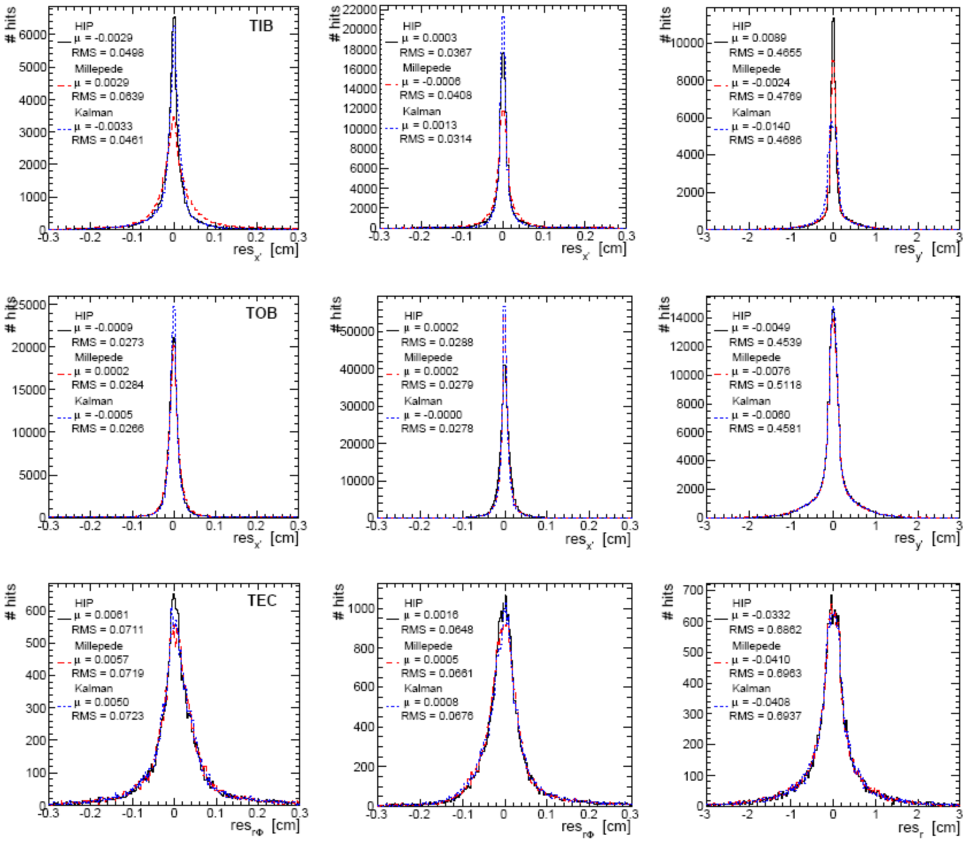


Figure 6: *Hit residuals for different geometries from three track-based algorithms: HIP (solid/black), Millepede (dashed/red), and Kalman (dotted/blue) based alignment. Three Tracker sub-detectors are shown in the top row (TIB), second row (TOB), and bottom row (TEC). The absolute local x'-residuals are shown for single-sided modules (left) and double-sided modules (middle), while local y'-residuals are shown for the double-sided modules only (right). For the endcap modules (TEC) transformation to the $r\phi$ and r residuals is made. The track fit is restricted to modules aligned by all three algorithms.*

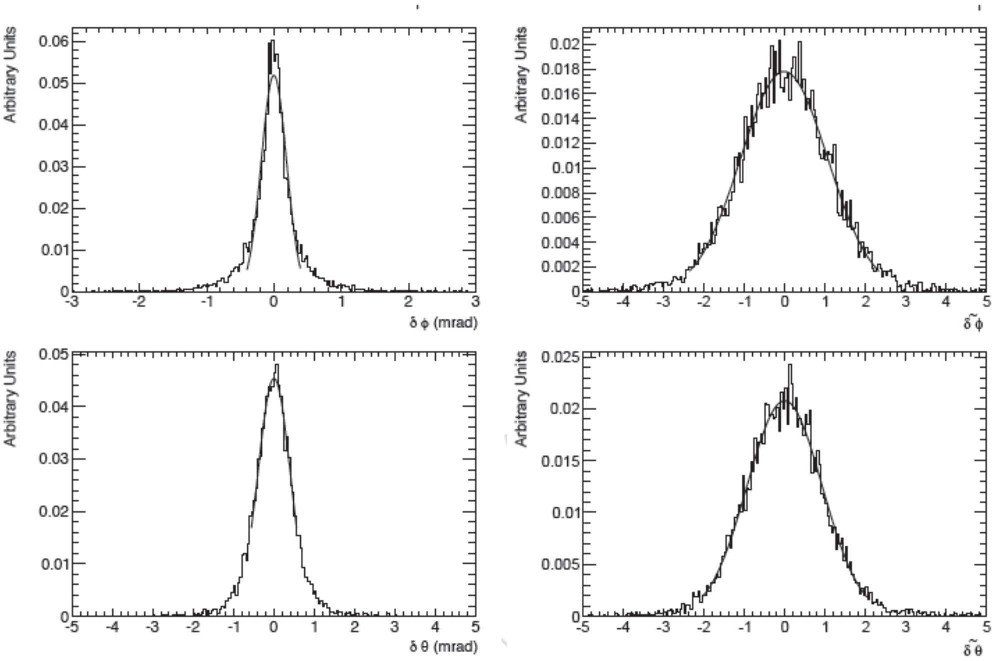


Figure 7: *Residual distribution (left) and pull distribution (right) of the azimuthal* ϕ *(top) and polar* θ *(down) angle.*

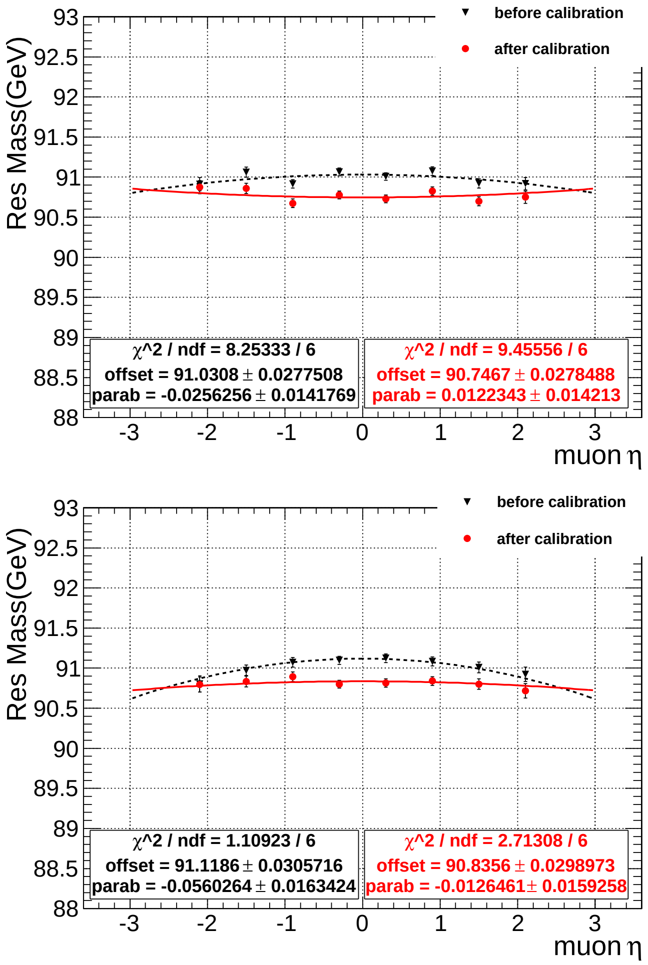


Figure 8: *Invariant di-muon mass for Z without (top) and with (bottom) QCD radiation computed in ideal detector conditions with global muons as a function of the pseudorapidity, before and after the calibration of the muon momentum scale.*

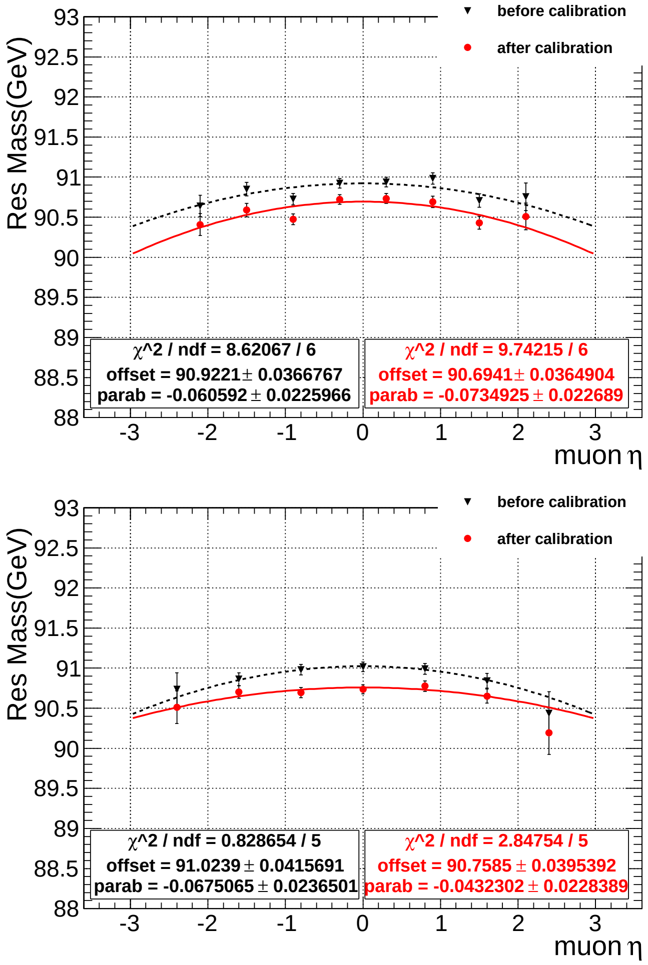


Figure 9: *Invariant di-muon mass for Z without (top) and with (bottom) QCD radiation computed on the fake data sample with global muons as a function of the pseudorapidity, before and after the calibration of the muon momentum scale.*

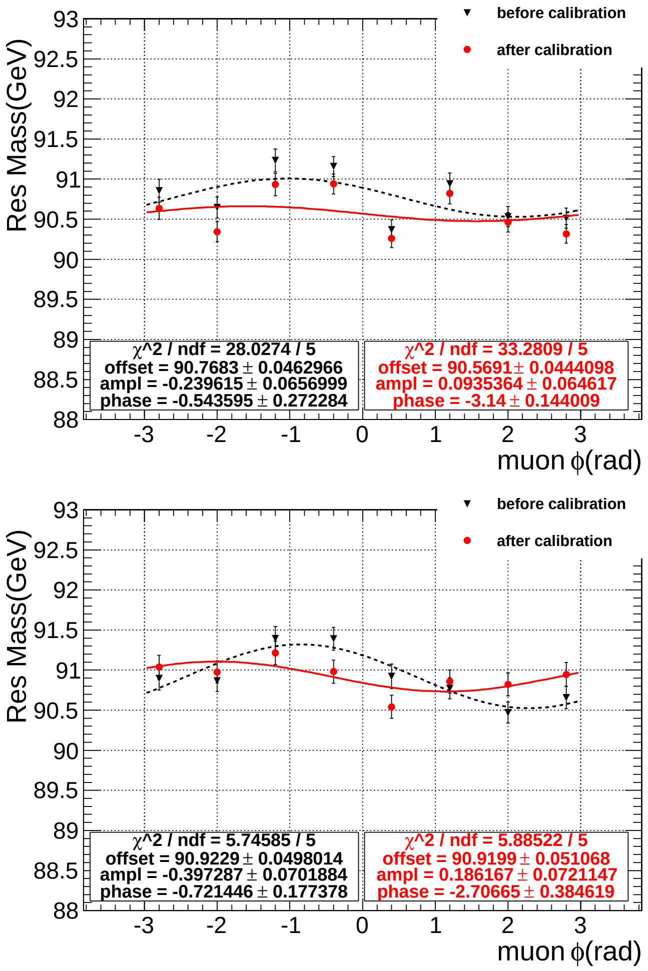


Figure 10: *Invariant di-muon mass for Z without (top) and with (bottom) QCD radiation computed on the fake data sample with global muons as a function of the polar angle, before and after the calibration of the muon momentum scale.*

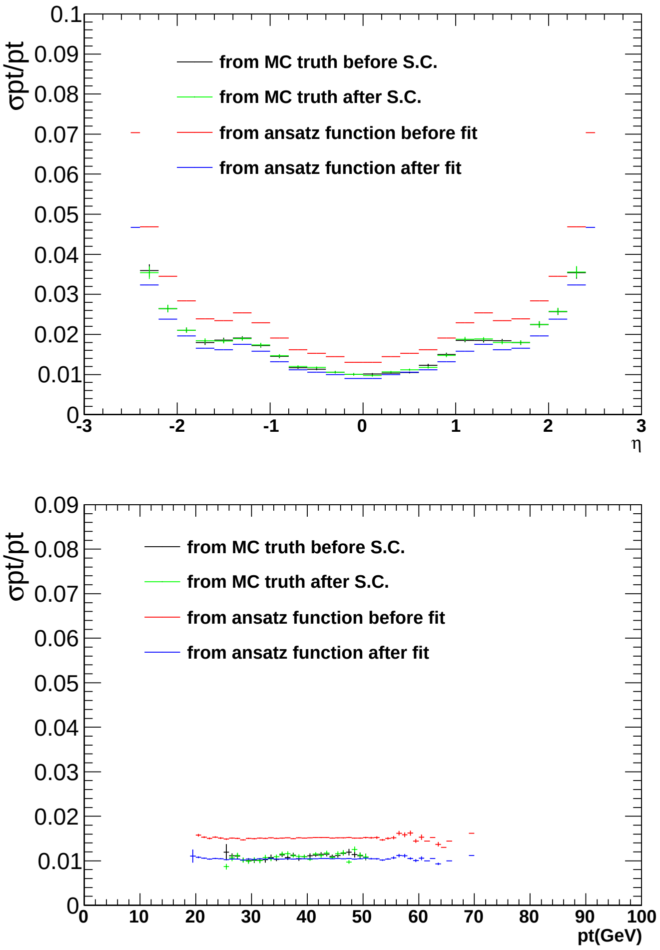


Figure 11: *Relative muon* p_T *resolution* (σ_pT/p_T) *as a function of the muon pseudorapidity (top) and the muon transverse momentum (bottom) for the ideal Monte Carlo data sample of Z without QCD radiation, computed using global muons before and after the scale correction. The comparison with relative resolution obtained from MC truth is shown.*

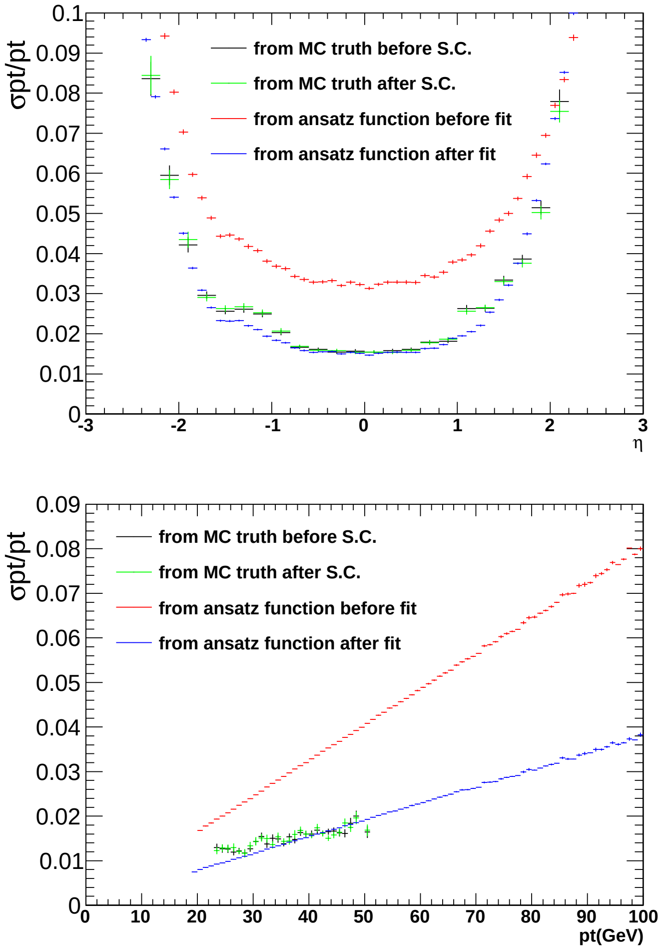


Figure 12: *Relative muon* p_T *resolution* (σ_pT/p_T) *as a function of the muon pseudorapidity (top) and the muon transverse momentum (bottom) for the fake data sample of Z with QCD radiation, computed using global muons before and after the scale correction. The comparison with relative resolution obtained from MC truth is shown.*

Ringraziamenti

É davvero lungo l'elenco di persone da ringraziare... non andró in ordine di importanza ma semplicemente per quello che sento.

Innanzi tutto, grazie a tutto il gruppo CMS Torino che mi ha accolta per questi tre anni, e mi ha permesso di passare cosí tanto tempo al CERN. Ho imparato davvero tanto in questi tre anni, e ve ne ringrazio molto! In particolare ovviamente un grande grazie a Marco, il mio Professore, per avermi sempre sostenuta all'interno del gruppo e per l'indipendenza che mi ha lasciato nelle mie scelte.

Un grazie enorme a Domenico, che mi ha salvata dai momenti piú difficili del mio dottorato, quelli in cui pensavo di mollare, quelli in cui non mi sentivo all'altezza, quelli in cui avevo ancora tante cose da imparare... Mi hai insegnato tutto quello che so sul software e sulle procedure di calibrazione. E tutto questo oltre a essere un amico, naturalmente!

Un grazie gigantesco anche a Sara, che ha sempre creduto in me e mi ha sempre sostenuta. Il tuo entusiasmo é contagioso e riempie di fiducia. E i tuoi consigli mirano sempre perfettamente il bersaglio! Sei una grande, e spero che tu rimanga al CERN a lungo!

Chiara, che dire... ho avuto la possibilitá di lavorare con te solo nella parte finale del mio dottorato... e devo dire che me ne dispiace molto, avrei voluto poter lavorare con te in maniera piú continuativa! Il tuo aiuto é stato prezioso, e la tua disponibilitá non é mai mancata, cosí come non sono mancati i tuoi incoraggiamenti. Grazie, davvero grazie di cuore! Spero che non ci perderemo di vista!

Posto importantissimo ha sicuramente la mia famiglia: sempre felicissima dei miei risultati, sempre entusiasta e quasi commossa per dove sono riuscita ad arrivare, dalla mia piccola cittá del Sud Italia che nulla offriva a un'adolescente assetata di scienza. Spero che i miei traguardi possano offrirvi soddisfazione e felicitá ancora per lungo tempo!

In ultimo ringrazio la persona piú importante: Pasquale. La cosa piú bella che mi sia mai capitata. Sei entrato nella mia vita solo da un anno, l'ultimo anno di questo percorso, il piú stressante. E sei stato solido come una roccia, sopportando anche i momenti in cui sono stata piú stressata, quelli in cui pensavo di perdere il controllo della situazione. Quelli in cui il tempo mi sembrava correre all'impazzata e mi lasciavo prendere dal panico. Mi hai reso tutto molto piú sereno e piú facile e incredibilmente piú bello! GRAZIE!!!

Bibliography

[1] W. Adam et al., *Performance studies of the CMS Strip Tracker before installation* Published on JINST:JINST 4 P06009, 2009.

[2] The CMS Collaboration, *Commissioning and Performance of the CMS Silicon Strip Tracker with Cosmic Ray Muons*, CMS PAPER CFT-09-002, arXiv:0911.4996, submitted to Jinst.

[3] F. J. Hasert et al., *Observation of neutrino-like interactions without muon or electron in the Gargamelle neutrino experiment*, Phys. Lett. **B46** (1973) 138.

[4] C. Rubbia, *Experimental Observation of the Intermediate Vector Bosons* W^{+}*,* W^{-}*, and* Z^{0}, Rev. Mod. Phys. **57** (1985) 699.

[5] J. F. Gunion, H. E. Haber, G. L. Kane et al., *The Higgs hunter's guide*, SCIPP-89/13.

[6] R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61. `hep-ex/0306033`.

[7] G. Ridolfi, *Search for the Higgs boson: Theoretical perspectives* (2001). `hep-ph/0106300`.

[8] G. Altarelli and G. Parisi, Nucl. Phys. **126** (1977) 297; V.N. Gribov and L.N. Lipatov, Sov. J. Nucl. Phys. **15** (1972) 438 and 675; Yu. L. Dokshitzer, Sov. Phys. JETP **46** (1977) 641.

[9] D. Acosta et al., Phys. Rev. D **70** (2005) 072002 [arXiv:hep-ex/0404004].

[10] S. Abdullin et al., CMS note 2006/067, available at http://cms.cern.ch/iCMS/

[11] V. Bĺuge et al., CMS note 2006/061, available at http://cms.cern.ch/iCMS/

[12] Workshop on Standard Model Physics (and more) at the LHC, CERN Yellow Report 2000-004.

[13] M. Davids et al., CMS note 2006/077, available at http://cms.cern.ch/iCMS/

[14] R. Chierici and A. Dierlamm, CMS note 2006/058, available at http://cms.cern.ch/iCMS/

[15] M. Spira, *HIGLU: A Program for the Calculation of the Total Higgs Production Cross Section at Hadron Colliders via Gluon Fusion including QCD Corrections* (1995). `hep-ph/9510347`.

[16] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012. hep-ph/0201195.

[17] D. Graudenz, M. Spira and P. M. Zerwas, *QCD corrections to Higgs boson production at proton proton colliders*, Phys. Rev. Lett. **70** (1993) 1372.

[18] M. Spira, A. Djouadi, D. Graudenz et al., *SUSY Higgs production at proton colliders*, Phys. Lett. **B318** (1993) 347.

[19] M. Spira, A. Djouadi, D. Graudenz et al., *Higgs boson production at the LHC*, Nucl. Phys. **B453** (1995) 17. `hep-ph/9504378`.

[20] R. V. Harlander and W. B. Kilgore, *Next-to-next-to-leading order Higgs production at hadron colliders*, Phys. Rev. Lett. **88** (2002) 201801. `hep-ph/0201206`.

[21] C. Anastasiou and K. Melnikov, *Higgs boson production at hadron colliders in NNLO QCD*, Nucl. Phys. **B646** (2002) 220. `hep-ph/0207004`.

[22] V. Ravindran, J. Smith and W. L. van Neerven, *NNLO corrections to the total cross section for Higgs boson production in hadron hadron collisions*, Nucl. Phys. **B665** (2003) 325. `hep-ph/0302135`.

[23] S. Catani, D. de Florian, M. Grazzini et al., *Soft-gluon resummation for Higgs boson production at hadron colliders*, JHEP **07** (2003) 028. `hep-ph/0306211`.

[24] A. Djouadi and P. Gambino, *Leading electroweak correction to Higgs boson production at proton colliders*, Phys. Rev. Lett. **73** (1994) 2528. `hep-ph/9406432`.

[25] A. Djouadi, P. Gambino and B. A. Kniehl, *Two-loop electroweak heavy fermion corrections to Higgs-boson production and decay*, Nucl. Phys. **B523** (1998) 17. `hep-ph/9712330`.

[26] K. G. Chetyrkin, B. A. Kniehl and M. Steinhauser, *Virtual top-quark effects on the $H \rightarrow b\bar{b}$ decay at next-to-leading order in QCD*, Phys. Rev. Lett. **78** (1997) 594. `hep-ph/9610456`.

[27] K. G. Chetyrkin, B. A. Kniehl and M. Steinhauser, *Three-loop $O(\alpha_s^2 G_F M_t^2)$ corrections to hadronic Higgs decays*, Nucl. Phys. **B490** (1997) 19. `hep-ph/9701277`.

[28] M. Spira, *QCD effects in Higgs physics*, Fortsch. Phys. **46** (1998) 203 - 284, arXiv:`hep-ph/9705337`.

[29] A. Djouadi, *The anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, arXiv:`hep-ph/0503172`.

[30] T. Han and S. Willenbrock, *QCD correction to the $pp \rightarrow WH$ and ZH total crosssections*, Phys. Lett. **B273** (1991) 167 - 172.

[31] O. Brein, A. Djouadi, and R. Harlander, *NNLO QCD corrections to the Higgs-strahlung processes at hadron colliders*, Phys. Lett. **B579** (2004) 149 - 156, arXiv:`hep-ph/0307206`.

[32] M. L. Ciccolini, S. Dittmaier, and M. Kramer, *Electroweak radiative corrections to associated WH and ZH production at hadron colliders*, Phys. Rev. **D68** (2003) 073003, arXiv:`hep-ph/0306234`.

[33] Z. Kunszt, *Associated Production of Heavy Higgs Boson with Top Quarks*, Nucl. Phys. **B247** (1984) 339.

[34] J. F. Gunion, *Associated top anti-top Higgs production as a large source of $W H$ events: Implications for Higgs detection in the lepton neutrino gamma gamma final state*, Phys. Lett. **B261** (1991) 510-517.

[35] W. Beenakker, S. Dittmaier, M. Kraemer, B. Pluemper, M. Spira, and P. Zerwas, *Higgs radiation off top quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **87** (2001) 201805, arXiv:`hep-ph/0107081`.

[36] S. Dawson, L. H. Orr, L. Reina, and D.Wackeroth, *Associated top quark Higgs boson production at the LHC*, Phys. Rev. **D67** (2003) 071503, arXiv:`hep-ph/0211438`.

[37] H.-Q. Zheng and D.-D. Wu, *First order QCD corrections to the decay of the Higgs boson into two photons*, Phys. Rev. **D42** (1990) 3760-3763.

[38] M. Steinhauser, *Corrections of $O(\alpha_s^2)$ to the decay of an intermediate-mass Higgs boson into two photons*, arXiv:`hep-ph/9612395`.

[39] A. Djouadi, P. Gambino, and B. A. Kniehl, *Two-loop electroweak heavy fermion corrections to Higgs-boson production and decay*, Nucl. Phys. **B523** (1998) 17-39, arXiv:`hep-ph/9712330`.

[40] U. Aglietti, R. Bonciani, G. Degrassi, and A. Vicini, *Two-loop light fermion contribution to Higgs production and decays*, Phys. Lett. **B595** (2004) 432-441, arXiv:`hep-ph/0404071`.

[41] G. Degrassi and F. Maltoni, *Two-loop electroweak corrections to the Higgs-boson decay $H \rightarrow \gamma\gamma$*, arXiv:`hep-ph/0504137`.

[42] J. Fleischer and F. Jegerlehner, *Radiative Corrections to Higgs Decays in the Extended Weinberg-Salam Model*, Phys. Rev. **D23** (1981) 2001-2026.

[43] D. Y. Bardin, B. M. Vilensky, and P. K. Khristova, *Calculation of the Higgs boson decay width into fermion pairs*, Sov. J. Nucl. Phys. **53** (1991) 152-158.

[44] B. A. Kniehl, *Radiative corrections for $H \rightarrow ZZ$ in the standard model*, Nucl. Phys. **B352** (1991) 1-26. doi:`10.1016/0550-3213(91)90126-I`.

[45] B. A. Kniehl, *Radiative corrections for $H \rightarrow W^+W^-$ in the standard model*, Nucl. Phys. **B357** (1991) 439-466. doi:`10.1016/0550-3213(91)90476-E`.

[46] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali, Phys. Lett. B, **429**, 263 (1998).

[47] L. Randall and R. Sundrum, Phys. Rev. Lett., **83**, 3370 (1999).

[48] L. Vacavant and I. Hinchliffe, J. Phys. G, **27**, 1839 (2001).

[49] B. C. Allanach et al., JHEP, 0009, 019 (2000).

[50] W.Witzeling and ATLAS Collaboration, *The ATLAS Detector: A Status Report* (ATL-CONF-99-006), CERN, Geneva (1999).

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

[52] S. Dasu et al., *CMS: The TriDAS project. Technical design report, vol. 1: The trigger systems*, CERN-LHCC-2000-038.

[53] The CMS collaboration, *CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger*, CERN-LHCC-2002-026.

[54] CMS Collaboration, *CMS: The Tracker Project Technical Design Report*, CERN/LHCC 98-06, CMS TDR 5, 15 April 1998,

CMS Collaboration, *Addendum to the CMS Tracker TDR*, CERN/LHCC 2000-016, CMS TDR 5 Addendum 1, 21 February 2000.

[55] M.Lozano, E.Cabruja, A. Collado, J. Santander and M. Ullan, *Nucl. Instrum. Meth.*, A473 (2001) 95-101, *Bump Bonding of Pixel System.*

[56] L.L. Jones, M.J. French, Q. Morrissey, A. Neviani, M. Raymond, G. Hall, P Moreira and G. Cervelli, CERN/LHCC 99-09 162-166, *The APV25 Deep Submicron Readout Chip for CMS Detector*,

L.L. Jones, *APV25-S1 User Guide Version 2.2*, 5th September 2001, available at http://www.te.rl.ac.uk/med.

[57] C. Civinini, *Proceedings of the International Conference on High Energy Physics*, hep2001, Budapest (Hungary) 12-18 July 2001, *The CMS Silicon strip sensors*

[58] CMS Collaboration, *CMS: The Trigger and Data Acquisition Project, Volume II: Data Acquisition and High-Level Trigger Technical Design Report*, CERN/LHCC 02-26, CMS TDR 6.2, 15 December 2002.

[59] S. Gadomski, G. Hall, T. Hogh, P. Jalocha, E. Nygard and P. Weilhammer, Nucl. Instrum. Meth., **A320** (1992) 217-227, CERN-PPE-92-024, Jan 1992, *The Deconvolution Method of Fast Pulse Shaping at Hadron Colliders.*

[60] R. Wunstorf, T. Rohe and A. Rolf, Nucl. Instrum. Meth., A388 (1997) 308-313, *Simulation of Irradiation-Induced Surface Effects in Silicon Detector.*

[61] S. Braibant, N. Demaria, L. Feld, A. Frey, A. Furtjes, W. Glessing, R. Hammarstrom, A. Honma, M. Mannelli, C. Mariotti, P. Mattig, E. Migliore, S. Piperov, O. Runolfsson, B. Schmitt and B. Surrow, CMS NOTE-2000/011, *Investigation of Design Parameters and Choice of Substrate Resistivity and Crystal Orientation for the CMS Silicon Microstrip Detector.*

[62] B.C. MacEvoy, G. Hall and K. Gill, Nucl. Instrum. Meth., A374 (1996) 12-26, *Defect Evolution in Irradiated Silicon Detector Materia.*

[63] H. Feick, E. Fretwurst, G. Lindstroem and M. Moll, Nucl. Instrum. Meth., A377 (1996) 217-223, *Long Term Damage Studies Using Silicon Detectors Fabricated from Different Starting Materials and Irradiated with Neutrons, Protons and Pions*;

J.A.J. Matthews, P. Berdusis, M. Frautschi, J. Schuler, H. Sadrozinski, K. O'Shaughnessy, L. Spiegel, A. Palounek, H. Ziock, N. Bacchetta, D. Bisello and A. Giraldo, Nucl. Instrum. Meth., A381 (1996) 338-348, *Bulk Radiation Damage in Silicon Detectors and Implications for LHC Experiments.*

[64] R. Wunstorf, PhD Thesis, *Systematische Untersuchungen zur Strahlenresistenz von Silizum-Detektoren fĺur die Verwendung in Hochenergiephysik-Experimenten*, University of Hamburg, October 1992.

[65] W. Adam et al. *The CMS tracker operation and performance at the Magnet Test and Cosmic Challenge* 2008. Published in JINST 3:P07006,2008.

[66] The CMS collaboration, *Detector Performance and Software Physics Technical Design Report, Volume I*, **CERN/LHCC 2006-001, CMS TDR 8.1.**; The CMS collaboration, *The CMS experiment at the CERN LHC*, JINST 3 (2008) S08004.

[67] The CMS Tracker collaboration, *CMS Tracker Alignment at the Integration Facility* - CMS Note 2009-002

[68] R. Fruhwirth, Application Of Kalman Filtering To Track And Vertex Fitting, Nucl. Instrum. Meth. A 262 (1987) 444.

[69] V. Karimaki, T. Lampen, F.-P. Schilling, *The HIP Algorithm for Track Based Alignment and its Application to the CMS Pixel Detector*, CMS NOTE-2006/018

[70] E. Widl, R. FrühwirthW. Adam, *A Kalman Filter for Track-based Alignment*, CMS NOTE-2006/022

[71] G. Flucke, P. Schleper, G. Steinbrück,M. Stoye, *A Study of Full Scale CMS Tracker Alignment using High Momentum Muons and Cosmics*, CMS NOTE-2008/008

[72] W. Adam et al., *Alignment of the CMS silicon strip tracker during stand-alone commissioning* Published on JINST: JINST 4 T07001, 2009

[73] W. Adam et al., *The CMS Detector Operation and Performance at the Cosmic Run at 666 Four Tesla*, **X** (2009). E. Widl, R.

[74] Fruhwirth and W. Adam, *A Kalman Filter for Track-based Alignment*, CMS-NOTE-2006-022.

[75] W. Adam et al., *Alignment of the CMS inner tracking system with cosmic ray particles*, 697 X (2009).

[76] R. Bainbridge et al., *Commissioning and Calibrating the CMS Silicon Strip Tracker*, 12th Workshop on Electronics For LHC and Future Experiments, Valencia (2006) 419-423.

[77] P. Biallass, T. Hebbeker, K. Hĺopfner, *Simulation of Cosmic Muons and Comparison with Data from the Cosmic Challenge using Drift Tube Chambers*, CMS Note 2007/024.

[78] M. Boezio et al., *Energy spectra of atmospheric muons measured with the CAPRICE98 balloon experiment* Phys. Rev. D **67**, 072003 (2003).

[79] The CMS Collaboration, *Detector Performance and Software Physics Technical Design Report*, Vol. 1, CERN/LHCC 2006/001, CMSTDR 8.1

[80] C. Leonidopoulus, *Physics and Data Quality Monitoring at CMS*, Proceedings of CHEP06, Mumbai, India (2006).

[81] C. Anastasiou et al., *High-precision QCD at hadron colliders: Electroweak gauge boson rapidity distributions at NNLO*, Phys. Rev. **D69** (2004) 094008. hep-ph/0312266.

[82] C. E. Gerber et al., *Tevatron-for-LHC Report: Top and Electroweak Physics* (2007). arXiv:0705.3251.

[83] C. Buttar et al., *Standard Model Handles and Candles Working Group: Tools and Jets Summary Report* (2008). arXiv:0803.0678.

[84] The CMS Collaboration, *Towards a measurement of the inclusive* $W \to \mu\nu$ *and* $Z \to \mu^+\mu^-$ *cross sections in pp collisions at* $\sqrt{s} = 14$ *TeV*, CMS Physics Analysis Summary 2007/002.

[85] N.Adam, V. Halyo, S. Yost *Updated theoretical uncertainties in elctroweak boson production cross sections for proposed CMS cuts at 10 TeV*

[86] T. Gleisberg et al., *Event generation with SHERPA 1.1*, (2008), 0811.4622.

[87] C.M. Carloni Calame, G. Montagna, O. Nicrosini and A. Vicini, *Precision electroweak calculation of the production of a high transverse-momentum lepton pair at hadron colliders*, JHEP 0710:109,2007 (arXiv:0710.1722 [hep-ph])

[88] S. Catani et al., *QCD Matrix Elements + Parton Showers*, JHEP 0111 (2001) 063, arXiv:hep-ph/0109231.

[89] F. Krauss, *Matrix Elements and Parton Showers in Hadronic Interactions*, JHEP 0208 (2002) 015, arXiv:hep-ph/0205283.

[90] A. SchÍalicke and F. Krauss, *Implementing the ME+PS merging algorithm* JHEP 0507 (2005) 018, arXiv:hep-ph/0503281.

[91] R. Hamberg, T. Matsuura and W.L. van Neerven, Nucl. Phys. **B345** (1990) 331; Nucl. Phys. B359 (1991) 343. W.L. van Neerven and E.B. Zijlstra, Nucl. Phys. B382 (1992) 11.

[92] U. Baur, S. Keller and D. Wackeroth, Phys. Rev. **D59** (1999) 013002.