Università degli Studi di Torino

Facoltà di Scienze Matematiche, Fisiche e Naturali
Dottorato di Ricerca in Fisica -XXII ciclo
FIS/01 - Fisica Sperimentale

Muon Detection and Reconstruction at CMS and their contribution to the Higgs Analysis Searches

Candidato
Dr.ssa Giorgia Mila

Tutore
Dr.ssa Chiara Mariotti

Coordinatore
Prof. Guido Boffetta

Ai torinesi di Rue de Lyon

Contents

Introduction **1**

1 Physics with the Large Hadron Collider (LHC) **5**
1.1 Why the LHC 5
1.1.1 LHC parameters 6
1.1.2 Phenomenology of the $p-p$ Collisions 11
1.1.3 LHC Experiments 13
1.2 CMS physics investigation 14
1.2.1 The SM Higgs search 17
1.2.1.1 Higgs production modes 17
1.2.1.2 Higgs decay channels and detection strategies 19
1.2.1.3 Higgs mass bounds 23

2 The Compact Muon Solenoid **27**
2.1 CMS Overview 27
2.2 The Inner Tracking System 30
2.2.1 The Pixel Detector 32
2.2.2 The silicon Microstrip Detector 33
2.3 The Calorimeters 34
2.3.1 The electromagnetic calorimeter 35
2.3.2 The Hadron Calorimeter 37
2.4 The Superconducting Magnet 39
2.5 The Muon System 41

2.5.1 The Drift Tube Chambers 42
2.5.2 The Cathode Strip Chambers 46
2.5.3 The Resistive Plate Chambers 47
2.6 Forward Detectors 48
2.6.1 CASTOR 48
2.6.2 Zero Degree Calorimeter (ZDC) 49

3 The Drift Tube calibration procedure 51
3.1 The DT detection principles 52
3.2 Calibration parameters 54
3.2.1 Calibration of the Time Pedestals 54
3.2.1.1 The inter-channel synchronization 56
3.2.1.2 The time pedestals computation 58
3.2.2 Calibration of the Drift Velocity 62
3.2.2.1 The Mean Timer method 63
3.2.2.2 The velocity drift computation 64
3.2.3 Analysis on the drift velocity 68
3.3 DT noise studies 71
3.4 The DT calibration workflow 75

4 Local Reconstruction in the Drift Tubes 81
4.1 Reconstruction of hits position within cells 82
4.1.1 The constant drift velocity method 83
4.1.2 Results with cosmic data 84
4.2 Segment Reconstruction 87
4.2.1 Segment building: the procedure 88
4.2.2 Results with cosmic data 90
4.3 The DT online Data Quality Monitoring 94
4.3.1 The CMS Data Quality Monitoring 94
4.3.2 The online DQM system 97
4.3.2.1 Data processing 97
4.3.2.2 Visualization 99
4.3.2.3 Operation 99
4.3.3 The online monitoring of the DT local reconstruction 99

5 The CMS Muon reconstruction 103
5.1 The Kalman Filter algorithm . . . 103
5.2 The muon spectrometer reconstruction . . . 105
5.2.1 The seed generator . . . 105
5.2.2 Pattern Recognition and Track Fitting . . . 106
5.3 Muon Reconstruction with the Full CMS Detector . . . 108
5.3.1 Track Reconstruction in the Tracker . . . 108
5.3.2 The Global Muon Reconstruction . . . 109
5.4 Muon Identification . . . 110
5.5 Muon Reconstruction Performance with Data from CRAFT 110
5.5.1 Efficiencies . . . 114
5.5.2 Momentum resolution . . . 116
5.5.3 Charge assignment . . . 118
5.5.4 Muon identification . . . 120
5.6 The Muon Reconstruction Offline Monitoring and Certification 123
5.6.1 The offline DQM system . . . 123
5.6.1.1 Data processing . . . 123
5.6.1.2 Visualization . . . 126
5.6.1.3 Operation . . . 127
5.6.2 The Muon Reconstruction Offline Monitoring . . . 127
5.6.2.1 The Run Products . . . 127
5.6.2.2 The harvesting step . . . 130
5.6.3 The Muon Reconstruction Certification . . . 131

6 The Higgs search through the H→ZZ→ 4μ channel 135
6.1 Monte Carlo simulation of the relevant physics processes . . 136
6.1.1 Signal: $H \to ZZ^{(*)} \to 4\ell$ . . . 136
6.1.2 Background: $t\bar{t}$ . . . 138
6.1.3 Background: $Zb\bar{b} \to 4\ell$ . . . 138
6.1.4 Background: $q\bar{q} \to ZZ^{(*)} \to 4\ell$ . . . 138
6.2 Trigger selection . . . 139
6.3 Lepton properties . . . 142
6.3.1 Muon Isolation parameters . . . 143
6.3.2 Muon measurement uncertainties . . . 143

6.4 Event Skimming and Pre-Selection 145
6.4.1 Event Skimming 146
6.4.2 Event Pre-selection 146
6.5 Lepton Discriminating Observables 150
6.5.1 Isolation 152
6.5.2 Impact Parameter 155
6.5.3 Kinematics 155
6.6 Results 156
6.6.1 Event selection 158
6.6.2 Analysis sensitivity to the Higgs signal 160
6.6.3 Combined Results 163
6.7 Control of Backgrounds from Data and Systematic Uncertainties 165
6.7.1 Measurement of the ZZ background from data . . . 168
6.7.2 Measurement of the $Zb\bar{b}$ background from data . . . 174
6.7.3 Experimental uncertainties 177

Summary **179**

Acknowledgement **181**

Introduction

The Standard Model (SM), the theory which describes three of the four known fundamental interactions between the elementary particles, has been experimentally tested with very high accuracy over a wide range of low-energy phenomena. However this successful result need to be investigated at high energies. In addition the SM still present some open questions which could be explained at an higher energy scale. The main weak spot of the theory is the mechanism which gives mass to all the particles. This mechanism is explained through the introduction of a massive scalar particle, the Higgs boson, which has not yet been observed. This led to the construction of the Large Hadron Collider (LHC), the most powerful machine for proton-proton collision ever built, at the European Laboratory for Particle Physics (CERN).

The LHC running conditions will reach in the next years unprecedented energy and luminosity, compared with those of previous accelerators, opening a new era in the experimental high energy physics. Four exepriments will collect data at the LHC: two of them are the general purpose experiments ATLAS and CMS, while the others, the ALICE and LHC-b detectors, are designed for studies on heavy ion collisions and b-physics, respectively. This experiments will have to cope with a very challenging physics environment: huge event rate, background cross sections higher of more than nine orders of magnitude with respect to the signal, large radiation dose and big amount of minimum bias events coming from proton-proton collisions. For this reason the hardware materials and software algorithms developed for the LHC detectors are at the forefront of the present technology and

production quality.

This thesis has been done within the Compact Muon Solenoid (CMS) experiment, which has been designed to provide excellent measurements of physics processes with high-energy muons in the final state. In particular, CMS is expected to detect with high efficiency and precision the Higgs decay into four muons, which is the so-called "golden channel" for the Higgs discovery.

My work focused on muons. It started from the analysis of the working conditions of the Drift Tubes (DT), which are detectors dedicated to the measurement of the muon three-dimentional track segments in the CMS spectrometer, and followed by the implementation of a dedicated software for the offline monitoring of the muon reconstruction inside the full CMS detector. Finally, I studied the experimental signature of the Higgs *golden channel* by simulating the total amount of data collected during the first LHC running period at 10 TeV.

The preliminary and fundamental step to detect the position of a charged particle passage inside a Drift Tube is to calibrate the signals timing. The procedure is described in detail in Chapter 3. This method requires a low electronic noise, since it could affect the signal arrival time distribution thus leading the calibration to fail. I studied the behaviour of these noise events in terms of rate and interval in time getting to the definition of a "noisy cell". I've also dedicated part of my work to the analysis of the noisy cells spatial distribution and their amount as a function of the magnetic field strength and the presence of other different sub-detectors in the data acquisition.

Chapter 4 describes the algorithm used to reconstruct the three dimensional segments inside each Drift Tube chamber. These segments are used for detector performance studies and they're given as input to the global CMS muon reconstruction. I've investigated the DT reconstruction procedure and developed a software dedicated to the online monitoring of the segment multiplicity, resolution and quality using the tools available in the central Data Quality Monitoring (DQM) framework of CMS.

To reconstruct muon tracks inside the complete CMS detector is necessary to collect hits from each subdetector crossed by the muon and a set

of other related information, *i.e.* associated energy deposits or muon identification parameters. This matching procedure, explained in Chapter 5, is done by the muon global reconstruction. After a detailed study to spotlight the most relevant properties of reconstructed muons, I've implemented a DQM code currently used for the offline monitoring of the muon reconstruction. This software first creates the main distributions of interest and then performs a list of dedicated tests. It eventually gives a definitive report on the muon reconstruction status, declaring the reconstructed muons to be good or not for physics studies.

All the detector performance studies have been done using cosmic data collected during the whole commissioning period (from 1997), while the Data Quality Monitoring software was mainly used and tested since 2008, when global CMS data acquisition runs started and daily online and offline shifts were organised by the central management. This led to a fast and fundamental DQM improvement thanks to the constant link between shifter reports and software optimisation by code developers. The results presented in this thesis are referred to the first high statistics cosmic ray data taking period with magnetic field on, the so-called Cosmic Run At Four Tesla (CRAFT) of 2008.

Finally in Chapter 6 the discovery potential of the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ processes is investigated. They are the main expected discovery channels at the LHC over a wide range of possible Higgs mass values, thanks to their very clear experimental signature. My work focused on the 4μ final state, whose analysis strategy rely mostly on the muon detection studies presented in the previous chapters. The CMS sensitivity for the observation of a Higgs boson has been studied in the mass range from 115 GeV/c^2 to 250 GeV/c^2 and the analysis was performed using a sequential set of cuts. Results are shown in the context of the startup luminosity (L = 1 fb^{-1}) and centre of mass energy of 10 TeV at the LHC.

Chapter 1

Physics with the Large Hadron Collider (LHC)

The Large Hadron Collider [1] is a particle accelerator built at CERN. It is contained in a circular tunnel, with a circumference of 27 kilometres, at a depth ranging from 50 to 175 metres underground. The accelerator can provide two colliding proton beams, with an energy of 7 TeV per proton, as well as heavy ions beams with an energy of $Z/A \cdot 7$ TeV per nucleon. The beams are injected in bunches separated in time by 25 ns. Its design luminosity is 10^{34} $cm^{-2}s^{-1}$ although during the first period of physics running the LHC should reach the peak of 2×10^{33} $cm^{-2}s^{-1}$. These two luminosity regimes are commonly called "Low luminosity" and "High luminosity", respectively. It is foreseen to be operating for at least 10 years.

1.1 Why the LHC

The Standard Model (SM) describes three fundamental interactions with two gauge theories: the theory of the electroweak interaction, that unifies the electromagnetic and the weak interactions, and the theory of strong interactions or Quantum Chromodynamics (QCD). The SM for weak and electromagnetic interactions is constructed on a gauge theory with four

gauge fields corresponding to massless bosons. Since only the photon is massless, whereas Ws and Z are massive, something has to happen in order to preserve the electroweak unification.

The masses of the gauge fields, as well as the fermion ones, in the SM are generated by spontaneous symmetry breaking. Electroweak symmetry breaking (EWSB) in the SM is described by the Higgs mechanism. Although a renormalizable theory which perfectly reproduces the low-energy phenomenology is then obtained, at present the Higgs mechanism is the weak spot of the theory, as it implies the rising of a new scalar particle, the Higgs boson, which has not yet been observed. If the Higgs boson does not exist, some processes could violate unitarity and then renormalizability could be lost. However, even the discovery of the Higgs boson could imply the necessity of new physics, depending on its mass and on the hypothetical discovery of new elementary particles at high energies.

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 some tens of fb and in the energy range from 100 GeV to 1-2 TeV.

1.1.1 LHC parameters

The idea behind the Large Hadron Collider is to reuse the existing LEP tunnel to install a new $p - p$ collider. Considerable financial savings are obtained from the fact that the tunnel and several infrastuctures (including pre-accelerators) already exist. However, the size of the tunnel limits the centre-of-mass energy to 14 TeV, since the beams must be bent by dipole magnets while maximum field is currently limited at about 8 T.
The machine parameters are shown in Table 1.1.

The choice of proton beams is dictated by the following considerations:

- The hadrons allow exploration of a wide range of energies with fixed-energy beams: they are the natural choice for a discovery machine.

Parameter		p-p	$^{208}Pb^{82+}$
Energy per nucleon (TeV)	E	7	2.76
Dipole field at 7 TeV (T)	B	8.33	8.33
Design Luminosity ($cm^{-2}s^{-1}$)	$\mathcal{L}$	10^{34}	10^{27}
Bunch time separation (ns)		25	100
No. of bunches	k_B	2808	592
No. particles per bunch	N_p	1.15×10^{11}	7.0×10^{7}
β-value at IP (m)	$\beta^\star$	0.55	0.5
RMS beam radiius at IP (μm)	$\sigma^\star$	16.7	15.9
Luminosity lifetime (h)	τ_L	15	6

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

The protons are not elementary particles and in hard collisions the interaction involves their constituents (quarks and gluons), which carry a non fixed fraction of the proton energy.

- The protons allow the accelerator to reach higher luminosities with respect to anti-protons machine, as their production and storage is easier.

- In a circular collider of radius R, the energy loss per turn due to synchrotron radiation is proportional to $(E/m)^4/R$, where E and m are respectively the energy and mass of the particles accelerated. Therefore, due to their higher mass, the usage of protons imply a smaller energy loss for synchrotron radiation with respect to the electrons.

Fig.1.1 shows the cross section of different processes as a function of the centre-of-mass energies for $p-p$ collision, and introduces to two fundamental considerations.

First, the increase of the Higgs cross section with the centre-of-mass energy, instead of the constant $p-p$ cross section, leads to use the highest possible proton beam energy. Moreover the fact that the Higgs cross section

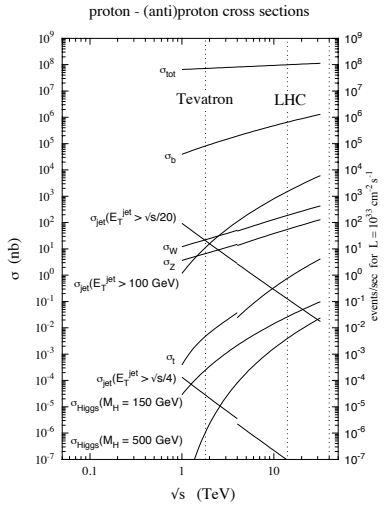


Figure 1.1: *Expected cross sections and event rates of several processes as a function of the centre-of-mass energy of proton-proton collisions [2].*

is many order of magnitude lower than the total cross section leads to operate at very high luminosity to accumulate statistics enough for physics analysis. The luminosity in fact represents the number of collisions per unit time and cross-sectional area of the beams. The relation between event rate R of a given process with cross section σ and the luminosity $\mathcal{L}$ is given by:

$$R = \mathcal{L}\sigma \tag{1.1}$$

The luminosity is specific to the collider parameters and does not depend on the process considered:

$$\mathcal{L} = \frac{\gamma f k_B N_p^2}{4\pi\epsilon_n\beta^\star}F, \tag{1.2}$$

where γ is the Lorentz factor, f is the crossing frequency, k_B is the number of bunches, N_p is the number of protons per bunch, ϵ_n is the normalised transverse emittance (with a design value of 3.75 μm), $\beta^\star$ is the betatron

function at the interaction point, and F is the reduction factor due to the crossing angle.

This design has the drawback that the total interaction rate becomes high and several interactions overlap in the same bunch crossing (*pile up*). As the total non-diffractive inelastic $p - p$ cross section predicted by PYTHIA [3] is 80 mb, on average 22 events will occur at every bunch crossing. With about 50 charged tracks per interaction, this pile-up poses several experimental problems, as discussed in Section 1.1.3.

The first beam was successfully steered around the full 27 kilometres of the LHC on 10 September 2008. To get beams around the ring in both directions on the first day exceeded all expectations, and the success continued through the night, with several hundred orbits being achieved. Then the LHC enjoyed a first week of commissioning with beam. Fig. 1.2 shows the LHC online monitoring while one of the two proton beams was circulating.

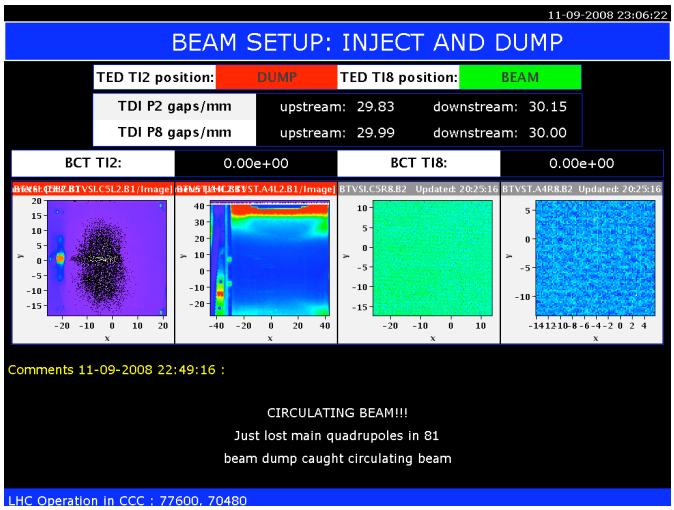


Figure 1.2: *Online monitoring of the LHC operation status during the first beam injections on September 2008.*

On Friday 19 September, during the commissioning (without beam) of the final LHC sector (sector 3-4) at high current, an incident occurred resulting in a large helium leak into the tunnel. The problem was caused by a faulty electrical connection between two magnets, which melted at high current leading to mechanical failure.

The CERN thus spent over a year repairing and consolidating the machine to ensure that such an incident cannot happen again. Recommissioning the LHC began in the summer 2009, and successive milestones have regularly been passed since then. The LHC reached its operating temperature of 1.9 Kelvin, or about -271 Celsius, on 8 October. Particles were injected on 23 October, but not circulated. A beam was steered through three octants of the machine on 7 November, and circulating beams were successfully injected on 20 November 2009. First collisions were seen by each of the fourth LHC experiments on 23 November 2009, when for the first time the two beams circulated simultaneously, as shown in Fig. 1.3.

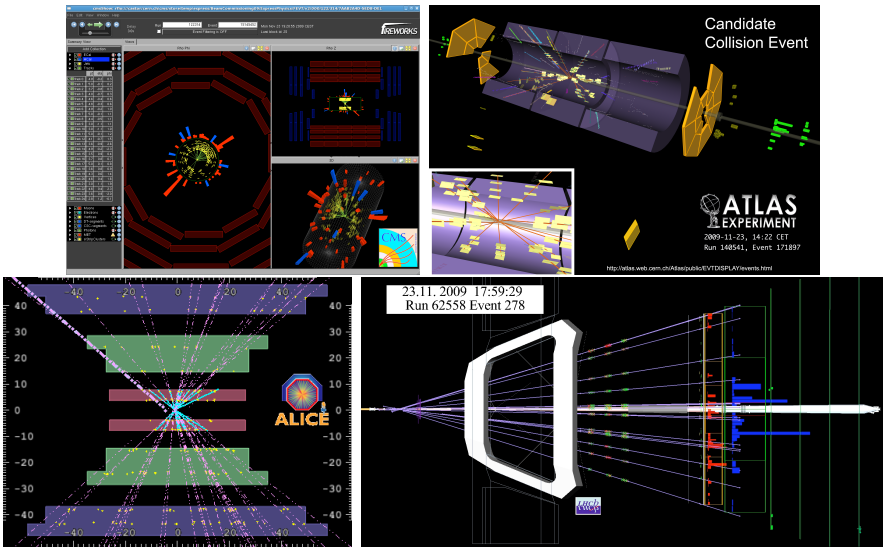


Figure 1.3: *First collisions seen by the four LHC experiments: CMS (top-left), ATLAS (top-right), ALICE (bottom-left), LHCb (bottom-right).*

On 30 November 2009 the LHC became the world's highest energy particle accelerator, having accelerated its twin beams of protons to an energy of 1.18 TeV in the early hours of the morning. This exceeded the previous world record of 0.98 TeV, which had been held by the US Fermi National Accelerator Laboratory's Tevatron collider since 2001, and marked another important milestone on the road to first physics at the LHC in 2010.

1.1.2 Phenomenology of the $p-p$ Collisions

One very remarkable aspect of LHC physics is the overwhelming background rate compared to the interesting physics processes: the Higgs production, for instance, has a cross section at least ten orders of magnitude smaller than the total inelastic cross section, as shown in Fig.1.1. In fact, the bulk of the events produced in $p-p$ collisions is either due to low-$\hat{p}_T$ scattering, where the protons collide at large distances, or to QCD high-$\hat{p}_T$ processes of the type:

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

All these events are collectively called "minimum bias" and in LHC studies are in general considered uninteresting since they constitute a background for other processes, where massive particles like the Higgs are created in the hard scattering. It should be noted that this classification is somewhat arbitrary because for example, this definition of minimum bias events includes $b\bar{b}$ production that is of interest for B-physics studies.

Another main characteristic of LHC is the span in energy of the initial state partons: in Fig.1.4 the CTEQ4M Parton Distribution Functions (PDFs) [4] at two different values of Q^2 are shown.

The fact that the two partons interact with unknown energies has two

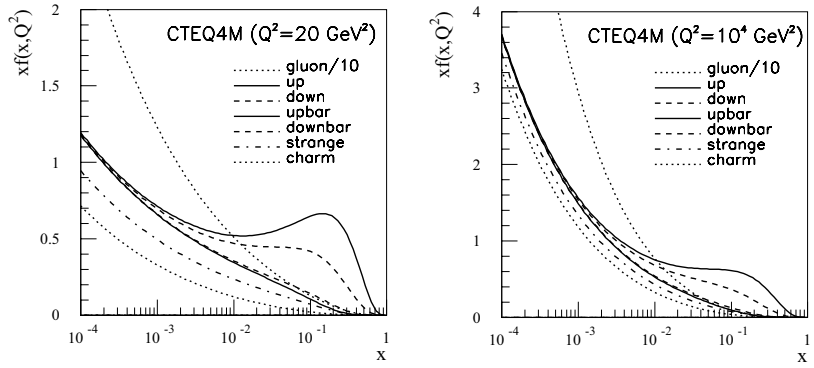


Figure 1.4: *Parton Density Functions for $Q^2 = 20\ GeV/c^2$ and $Q^2 = 10^4\ GeV/c^2$.*

fundamental consequences. First of all the total energy of an event is unknown, 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. It is therefore very useful to use experimental quantities that are invariant under such boosts as the transverse momentum p_T or the *rapidity (y)*. The rapidity, choosing the beam direction as z axis, is defined as:

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

and it is often used to describe angular distributions because it is additive under boosts along the z direction, therefore the shape of $dN{=}dy$ distributions is invariant under such boosts. For 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. The pseudorapidity can be reconstructed just from the measurement of the θ angle and can be also used for particles for which the mass and momentum are not measured.

1.1.3 LHC Experiments

LHC detectors will operate in a very difficult environment: the high bunch crossing frequency, the high event rate and the pile-up of several interactions in the same bunch crossing impose strict requirements on the detector design. To cope with a bunch crossing rate of 25 ns and a pile-up of about 20 interactions per bunch crossing, the detectors shall have a very fast time response and the readout electronics shall also be very fast. Due to the presence of pile-up, high granularity and a large number of electronic channels are required to avoid the overlap of particles in the same sensitive elements. LHC detectors will also have to withstand an extremely high radiation dose and special radiation-hard electronics must be used. Additional requirements apply to the on-line trigger selection, that has to deal with a background rate several orders of magnitude higher than the signal rate.

Four experiments are installed at the LHC. Two of them are devoted to specific topics: ALICE [5] to heavy ions and LHC-b [6] to b-physics. The other two are the general-purpose experiments ATLAS [7] and CMS [8]. Their design differs significantly, since two very different solutions were chosen for the magnetic field configuration: CMS uses a solenoidal field generated by a big superconductiong solenoid, while ATLAS uses a toroidal field produced by three sets of air-core toroids complemented by a small inner solenoid. A detector based on a toroidal magnet has the advantage that the track p_T resolution is constant as a function of the pseudorapidity. A very large air-core toroid allows a good momentum resolution even without the aid of the inner tracker; however, it requires very precise detectors with excellent alignment. An iron-core solenoid, on the other hand, can generate a very intense field. The resulting system is very compact and allows calorimeters to be installed inside the magnet, improving the detection and

energy measurement of electrons and photons. Precise tracking exploits both the constant field within the magnet and the field inside the return yoke. Moreover, tracks exiting the yoke point back to the interaction point, a property that can be used for track reconstruction. Multiple scattering within the yoke, however, degrades the resolution of the muon system.

Schematic pictures of CMS and ATLAS are shown in Fig.1.5 and Fig.1.6, respectively. The CMS experiment is described in details in Chapter 2.

1.2 CMS physics investigation

The CMS physics program covers a wide range of precision electroweak measurements and searches for new particles. It should be noted that the precision required will not be trivial to achieve being at a hadron collider where the initial state of the parton-parton collision is not well known and where the final state is complicated by the presence of many other particles produced. However, thanks to the high statistics available, for most measurements the statistical uncertainty will be very small. Moreover high statistics control samples will allow a good understanding of the detector response, thus reducing the systematic uncertainty.

The physics investigation of the CMS experiment can be summarised on three main items:

- study the mechanism that breaks the symmetry of the SM Lagrangian giving rise to the particle masses. Within the SM this means to search for the Higgs boson from $m_H = 100$ GeV to $m_H = 1$ TeV. If the Higgs is found, understand if it has the same properties predicted by the SM; if the Higgs is not found, look for alternative models.

- search for new physics, especially if the Higgs is not found. Concerning supersymmetry, all the s-particles with mass $m_{\tilde{s}} \leq 3$ TeV will be accessible. For exotic models (like lepto-quark, technicolor, new strong interaction, new lepton families, additional bosons, extra-dimensions) the mass reach is 5 TeV.

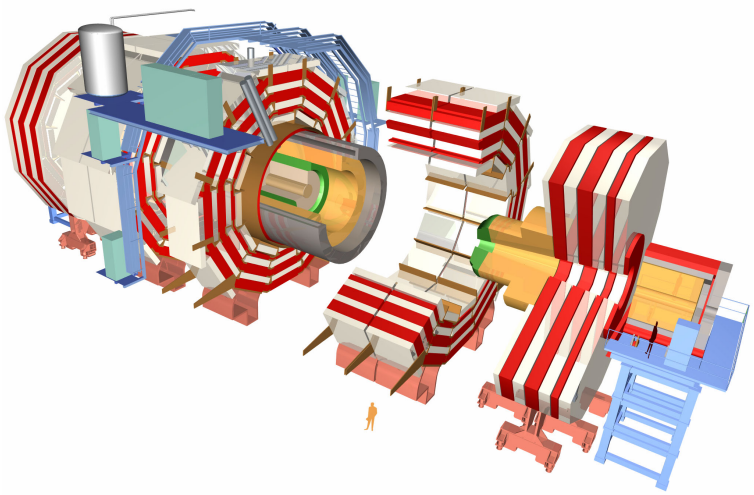


Figure 1.5: *The CMS detector*

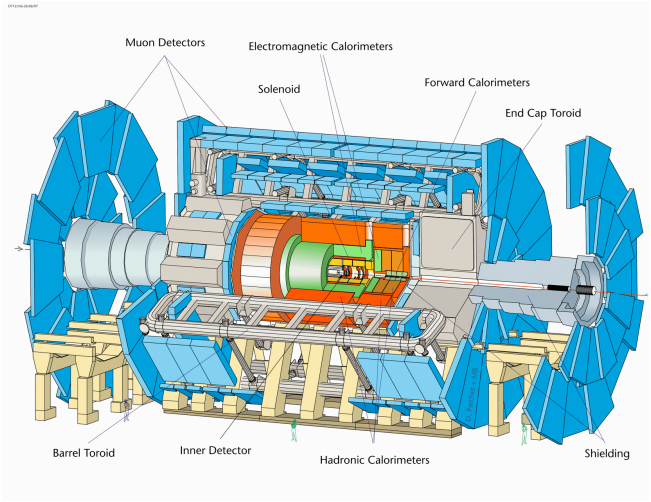


Figure 1.6: *The ATLAS detector*

- perform precision measurements in the electroweak sector (m_W, m_{top}, triple gauge couplings, $sin^2\theta_W$), in QCD, and in the CP violation and B physics sector. Concerning the precision electroweak measurements, the top mass and the W mass should be measured with a relative precision given by:

$$\Delta m_W = 0.7 \times 10^{-2} \Delta m_{top} \tag{1.5}$$

  in order to have a comparable impact in the determination of the Higgs mass from the fit of the electroweak observable. The target precision on these quantities is expected to be $\Delta\ m_W \leq 15$ MeV and $\Delta\ m_{top} \leq 2$ GeV.

The first years will be devoted to the Standard Model "re-discovery", a series of standard processes will be searched to function as *standard candles*: Drell- Yan with the Z peak and quarkonia low mass resonances, W production and $t\bar{t}$ production. The analysis of these processes will allow to calibrate the detectors: the lepton energy scale can be kept under control exploiting the di-lepton mass resonances (Z, J/ψ, and Υ), the missing transverse energy can be calibrated with the W transverse mass, finally with the huge rate of $t\bar{t}$ we can calibrate jets and measure the b-tagging efficiency. Moreover some processes will be very useful at the beginning to tune the Monte Carlo SM description. 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. Analogously, the measure of the W^+/W^- asymmetry will allow to increase the precision on the PDF. 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. More detailed information about the CMS physics program can be found elsewhere [9, 10]. As an introduction to the analysis presented in Chapter 6, a brief review of the Higgs physics at LHC is here reported.

1.2.1 The SM Higgs search

In the Standard Model (SM) the elementary particles acquire their mass through the Higgs mechanism. This mechanism foreseens the existence of the Higgs boson, a scalar particle which couples to massive particles. Its mass is the only yet unknown parameter of the SM. In the following the main Higgs production and decay channels are described focusing in particular on their experimental signature. Finally, the Higgs boson total decay width and the theoretical and experimental bounds to the Higgs mass are shortly illustrated.

1.2.1.1 Higgs production modes

Fig.1.7 shows the tree level diagrams of the four main Higgs production channels in $p-p$ collisions and Fig.1.8 illustrates their cross sections, as a function of the Higgs mass, for a center of mass energy of 14 TeV.

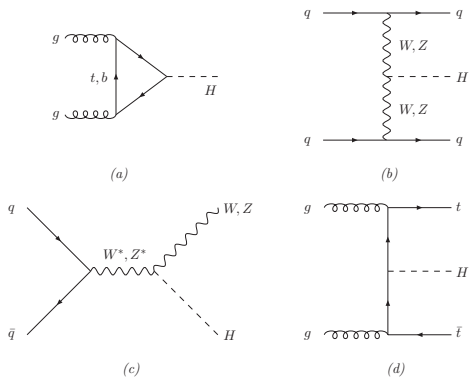


Figure 1.7: *Higgs boson production mechanisms at tree level in proton-proton collisions: (a) gluon-gluon fusion; (b) Vector Boson Fusion, (c) W, Z associated production (or Higgsstrahlung); (d) $t\bar{t}$ associated production.*

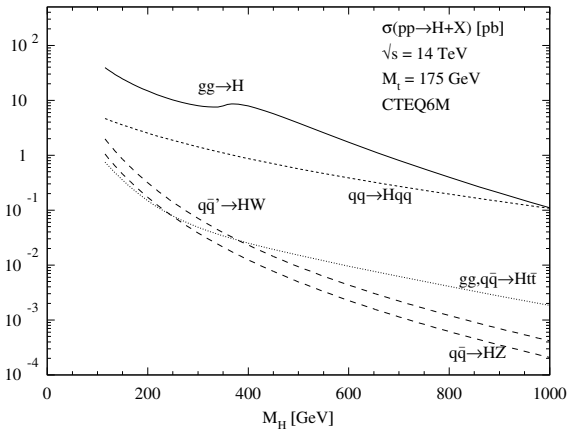


Figure 1.8: *Higgs boson production cross sections at $\sqrt{s}$ = 14 TeV as a function of the Higgs boson mass. The cross sections are calculated using HIGLU [11]; they contain higher order corrections and the CTEQ6m [12] p.d.f. has been adopted.*

The **gluon-gluon fusion** is the dominant process over the whole mass spectrum. The process is shown in Fig.1.7(a), with a t quark-loop as the main contribution. Its lowest order cross section suffers of high order QCD corrections. The increase in cross section from higher order diagrams is conventionally defined as the K-factor:

$$K = \frac{\sigma_{NLO}}{\sigma_{LO}} \tag{1.6}$$

where LO (NLO) refer to leading (next-to-leading) order results. The NLO QCD corrections to the top and bottom quark loops [13, 14, 15] results in a 50 to 100% increase of the cross section. Moreover large uncertainties due to the gluon structure functions affect the cross section computation. Finally, it should be noted that the production of the Higgs boson through gluon fusion is sensitive to a fourth generation of quarks. Because the Higgs boson couples proportionally to the fermion mass, including a fourth generation of very heavy quarks, its cross section will increase more than twice.

The **Vector Boson Fusion** (VBF) (Fig.1.7(b)) cross section is about one order of magnitude lower than the gg fusion for a large range of the Higgs masses. The two processes become comparable for high values ($\mathcal{O}$(1 TeV)) of the Higgs mass. This process has a well known next-to-leading-order cross section and small higher order QCD corrections. Moreover, this channel is very interesting because of its clear experimental signature: the presence of two spectator jets with high invariant mass in the forward region provides a powerful tool to tag the signal events and discriminate the backgrounds, thus improving the signal to background ratio, despite its low cross section.

The remaining production processes, the so-called **association production channels**, have very small cross sections, orders of magnitude lower than those of gg and VBF. They will be used for the Higgs discovery in association with particular Higgs decay modes to exploit final states with a clear signature.

In the *Higgsstrahlung* process (Fig.1.7(c)), the Higgs boson is produced in association with a W^+ or Z boson, which can be used to tag the event. The QCD corrections are quite large and the next-to-leading order cross section is increased by a factor of $1.2 \div 1.4$ relative to the LO one.

The last process, illustrated in Fig.1.7(d), is the *associated production of the Higgs boson with a $t\bar{t}$ pair*, which can provide a good experimental signature. The higher order corrections increase the cross section by a 1.2 factor.

1.2.1.2 Higgs decay channels and detection strategies

The branching ratio of all the possible Higgs decay channels as a function of the Higgs mass is shown in Fig.1.9. For Higgs masses up to $150 \sim \mathrm{GeV/c^2}$ fermionic decay modes are dominating. When the decay channels into vector boson pairs open up, they quickly dominate. At high masses (above $350 \sim \mathrm{GeV/c^2}$) also $t\bar{t}$ pairs can be produced. Depending on the Higgs mass, different strategies have been developed for its search. The experimental techniques and the expected sensitivity are discussed in the following.

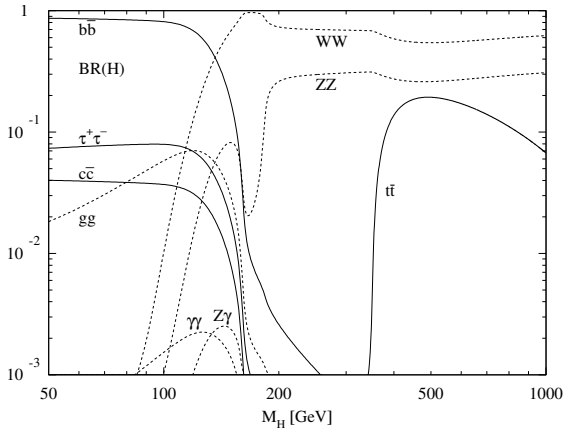


Figure 1.9: *Branching ratios for different Higgs boson decay channels as a function of the Higgs boson mass. They are calculated with the program HDECAY [16] which includes the dominant higher order corrections to the decay width.*

Low mass region

The most promising discovery channels for $M_H < 130 \sim \mathrm{GeV/c^2}$ are the decay modes into a pair of photons or τ leptons thanks to their clear signature.

The first process suffers of a very high background coming from Drell-Yan e^+e^-, $pp \to \gamma\gamma$, $pp \to jets + \gamma$, $pp \to jets$ where one or more jets are misidentified as γ. The suppression of the last two contributions will require a good understanding of the performance of the electromagnetic calorimeter and a reliable modeling of the amount of material in front of it.

The analysis on the $H \to \tau\tau$ decay mode focus on the VBF production channel because of its higher signal over background ratio. The most promising final state is the one with one τ decaying into leptons and the other into hadrons. The irreducible backgrounds to this process are the QCD and EW production of two τ leptons from Z or $\gamma^\star$ with associated jets.

Contributions also come from W+multi-jets production and $t\bar{t}$ events in which one of the jets can be misidentified as a τ jet. The signature is characterized by the hadronically decaying τ (associated to a little (ΔR=0.4) isolated jet), the leptonically decaying τ (identified from the electron or the muon with highest transverse momentum $p_T > 15 \sim \mathrm{GeV/c^2}$) and the two quarks emitting the bosons in the VBF process which have a high energy and rapidity gap.

The high branching ratio of the Higgs boson into a pair of b quarks can only be exploited in the study of the Higgs production via $t\bar{t}$ fusion. The most promising final states have at least one of the two t quark decaying leptonically thanks to the clear signature offered by the presence of at least one high p_t lepton from one of the two W, missing energy and 4 b-tagged jets (of which two from the Higgs). The four jets are the responsible of a very high background, mostly composed by $ttbb$, Zbb, $tt + Njets$ and multi-jets QCD events and are the main sources of uncertainty. A pioneer novel study [17] has obtained good results on the signal over background ratio by reconstructing the $H \to b\bar{b}$ decay (produced through VBF) asking the presence of a central high p_t photon in the final state.

High mass region

This region corresponds to values of the Higgs boson mass above the threshold of $2M_W$, where the Higgs analysis are focused on the Higgs decays into a couple of vector bosons.

The main channels of interest are those where the two vector bosons decay leptonically. The clear experimental signature of these events compensates for their low branching ratio, which is about one order of magnitude lower than the hadronic ones.

In the $H \to WW \to l\nu l\nu$ channel it is not possible to reconstruct the H invariant mass due to the presence of the two neutrinos. Since the signal selection can not exploit this variable, other techniques must be used for the discrimination and a good control of the background shape is mandatory. The final state presents two isolated high p_T leptons pointing to the primary vertex, high missing energy and no hadronic activity. The signal selection

relies mainly on the request of a central jet veto, high missing energy and of a small angle between the two leptons due to the V-A structure of the weak interaction.

The $H \rightarrow ZZ \rightarrow 4l$ (l = muons or electrons) channels are the "golden channels" for the Higgs discovery. The main backgrounds are: $t\bar{t}$, $Zb\bar{b}$ and the irreducible $ZZ^{\star}/\gamma^{\star}$. The trigger and the offline cuts rely on the presence of isolated charged leptons coming from the primary vertex, with high transverse momentum. The instrumental backgrounds become negligible with the request of lepton isolation and using different cuts on the sorted lepton transverse momenta. The irreducible background can be suppressed applying cuts on angular variables. The main sources of systematic uncertainties come from the choice of the PDF and the QCD scale, the NLO versus the LO dynamics, the isolation cut and its efficiency, the electron reconstruction efficiency, the energy and the momentum scale and the charge identification.

As discussed above, the VBF production channel becomes important in the very high mass region thanks to its clear experimental signature given by the two spectator jets and the Higgs decay products, which allows a good rejection of dominant background coming from $V + njets$, $VV + njets$ and $t\bar{t}$ production. These jets are well separated in pseudorapidity and have a very high invariant mass.

Moreover, the Vector Boson Fusion cross section (with or without a production of an Higgs particle) is an extremely interesting process because the cross section $\sigma(pp \rightarrow VVjj)$ and the polarization of the VV pair depend sensitively on the presence or absence of a light Higgs in the physical spectrum. If a massive Higgs boson exists, a resonance will be observed in the VV invariant mass spectrum in correspondence of the Higgs mass. In the absence of the Higgs particle the SM predicts that the scattering amplitude of longitudinally polarized vector bosons grows linearly with s ad violates unitariety at about $1 - 1.5$ TeV. This means that the measurement of the cross section at large $M(VV)$ could provide information on the existence of the Higgs boson independently of its direct observation.

It should be also noted that the Higgs width becomes very broad for masses above 500 GeV/c^2, as shown in Fig.1.10, so that the reconstruction

of its mass peak becomes difficult. The highest limit of detection is $M_H \sim$ 1 TeV, where the Higgs resonance width becomes larger than its own mass. In this mass range the Higgs boson cannot be properly considered as a particle any more. In addition, if the Higgs boson mass is above 1 TeV, the SM predictions violate unitarity. All these considerations suggest the TeV as a limit to the Higgs boson mass: at the TeV scale at least, the Higgs boson must be observed, or new physics must emerge.

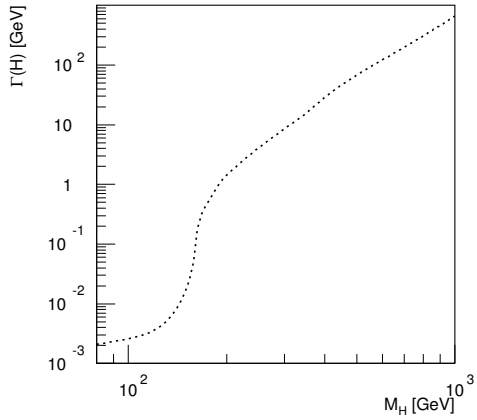


Figure 1.10: *Higgs total decay width as a function of the Higgs boson mass.*

1.2.1.3 Higgs mass bounds

Theoretical Bounds

The mass of the Higgs boson is not predicted by the SM; however arguments of self-consistency of the theory can be used to yield constraints on its value.

An upper bound on m_H can be found requiring the Higgs self coupling λ to remain finite in its running up to an energy scale Λ (*triviality*). In fact λ

rises indefinitely with energy and the theory could eventually become non-perturbative. On the other hand, the requirement that, when radiative corrections are included, the minimum of the Higgs field potential is an absolute minimum up to the scale Λ leads to a *vacuum stability* condition which limits m_H from below.

In both cases, the parameter Λ represents the scale up to which the Standard Model is taken to be valid. The theoretical bounds on m_H as a function of Λ are shown in Fig.1.11.

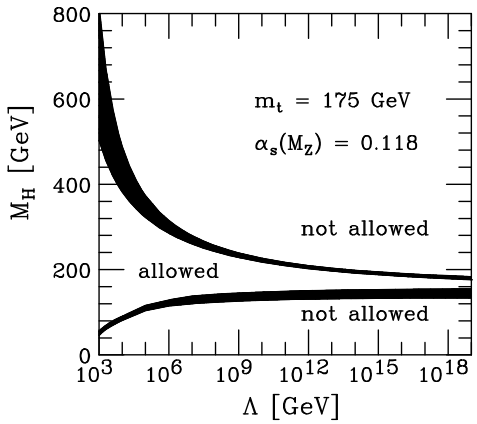


Figure 1.11: *Theoretical bounds on the Higgs mass as a function of the energy scale Λ up to which the Standard Model is valid [18].*

Only the range 130-190 GeV/c^2 allows the SM o remain valid up to the Plack scale ($\Lambda = 10^{19}$ GeV). Assuming the SM to be valid up to $\Lambda \sim 1$ TeV, the Higgs mass can be up to 700 GeV/c^2. In any case, on the basis of these results, the guideline for the design of present and future colliders should be the search of the Higgs boson up to masses of ≈ 1 TeV.

Experimental Bounds

Experimental bounds on m_H are provided by measurements at LEP, SLC and Tevatron. Direct searches from LEP [19] excluded the possibility of a Higgs boson with a mass below 114.4 GeV/c^2 at 95% confidence level. Indirect bounds can be obtained from precision electroweak measurements; observable like the top and $W^{\pm}$ masses and Z decay parameter are logarithmically sensitive to m_H through radiative corrections. These data can therefore be fitted taking the Higgs mass as free parameter. The results of this procedure are summarised in Fig.1.12. The plot shows the value of $\Delta\chi^2 = \chi^2 - \chi^2_{min}$ as a function of the Higgs mass. The fit privileges low values of the Higgs mass and an upper limit of 160 GeV/c^2 can be set at 95% confidence level.

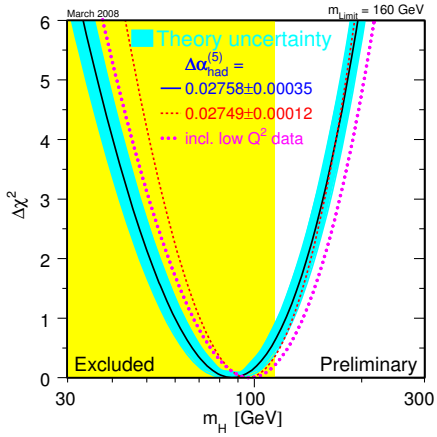


Figure 1.12: *Global fit to electroweak precision measurements done by the LEP Electroweak Working Group for Winter 2008*

Finally, it should be noted that the Tevatron CDF and D0 experiments are actually working on Higgs direct searches. Because of the low statistics, their combined analysis are focused not on the mass peak search but on the

determination of mass energy ranges where the presence of the Higgs boson can be excluded. The last results, presented on November 2009, exclude the existence of the Higgs boson in the region between 163 and 166 GeV/c^2 at 95% of confidence level, as shown in Fig.1.13.

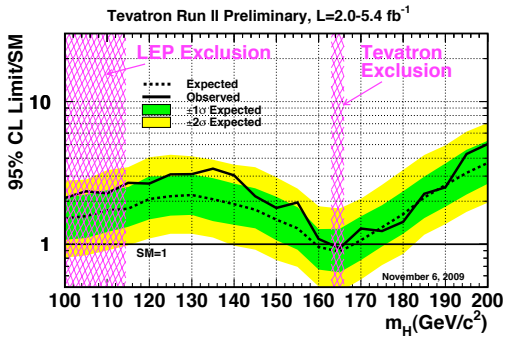


Figure 1.13: *95% confidence level of the combined Tevatron analysis on the Higgs searches as a function of the Higgs mass.*

Chapter 2

The Compact Muon Solenoid

2.1 CMS Overview

In order to fulfill its physics goals the CMS detector must be able to identify muons, electrons and photons and measure their properties with high precision. Jets, made up of particles closely correlated in space, also need to be measured. They arise through the hadronization of scattered quarks or gluons in a collision. For several new particle searches it is also important to determine if a jet originated from a beauty-quark or not. Neutrinos and other very weakly interacting particles escape without leaving signals. They can only be measured indirectly, through the determination of missing transverse energy, which requires a hermetic detector. Therefore the detection systems of CMS must cover as much of the solid angle as possible. For this purpose and to enhance the physics search capabilities, new forward detectors have been added to the original CMS design. The measurement technique of missing transverse energy relies on the fact that the sum of transverse energy vectors to all calorimeter cells should vanish due to conservation laws if no weakly interacting particles were produced in the collision. If the sum is non-zero, it corresponds to the negative of the

resulting transverse energy vector of one or more of these particles. Finally, to the CMS detector is required not only to identify particles and to determine their parameters but to assign them to the beam crossing from which they originated.

The final CMS layout is illustrated in Fig.2.1. The detector consists of a cylindrical barrel closed by two endcap disks. The overall length is 21.6 m, the diameter 14.6 m and the total weight about 12500 tons. Being the interaction point at the center of CMS, subdetectors have been arranged in a onion-like structure around it.

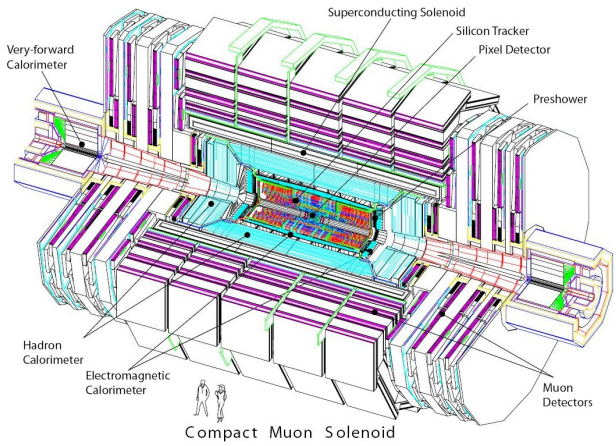


Figure 2.1: *The CMS detector.*

The CMS coordinate system is a right-handed frame where the x axis points to the centre of the LHC ring, the y axis points upward and the z axis is parallel to the beam.

As discussed in Sec.1.1.3, the subdetectors chose was dictated by the shape of the central solenoid [20] which has an inner radius of 2.95 m and a length of 13 m. It is a superconducting device cooled with liquid helium, producing a maximum magnetic field of 3.8 T. The inner tracking detectors and calorimeters are located inside the solenoid. The innermost detector

around the beam pipe (radius 43 mm) is a silicon pixel vertex detector [21] with three layers in the barrel and two disks in each endcap. This detector allows to reconstruct the exact positions of the interaction vertices, which is essential for b-tagging (identifying interactions that resulted in B mesons). Outside the pixel detector, a silicon strip detector extends up to a radius of 1.2 m. Together with the pixel detector it is used for the reconstruction of charged tracks. An electromagnetic calorimeter (ECAL) [22] is located outside the tracker with a coverage of up to $|\eta| < 3$. It is a homogeneous device consisting of a large number of scintillating crystals that are read out using avalanche photo diodes or vacuum photo-triodes. Its purpose is the precise measurement of the energy deposit of electrons and photons. In the endcap region a preshower detector is placed in front of the ECAL in order to improve spatial resolution and to provide pion/photon separation. It consists of two thin lead converters followed by silicon strip detector planes. The outermost detector inside the solenoid is the hadronic calorimeter (HCAL) [23] which provides the same pseudo-rapidity coverage as the ECAL. It is a sampling calorimeter consisting of copper absorber plates interleaved with scintillator sheets. Its purpose is the reconstruction of jets as well as the measurement of total and missing transverse energy. The very forward hadronic calorimeter (HF) extends the coverage of the HCAL up to $|\eta| < 5.3$ and enhances the hermeticity of the detector. It is placed around the beam pipe, outside the magnet and muon system. The muon system is embedded in the iron return yoke of the magnet, outside the coil. Three different technologies are used: drift tubes in the barrel up to $|\eta| < 1.2$, cathode strip chambers in the endcaps up to $|\eta| < 2.4$ and resistive plate chambers in the whole detector up to $|\eta| < 2.1$. All three detectors are used in the trigger. DT and CSC are also used in the reconstruction of muons coming from the interaction vertex. The forward region is instrumented with more detectors, which will be described in Section 2.6.

A transverse section of CMS is shown in Fig.2.2 . The signals which different kinds of particles leave in the subdetectors are also shown. Muons easily traverse the entire detector, leaving tracks in the inner tracker and in the muon chambers, whilst depositing almost no energy in the calorimeters.

Electrons and photons produce electromagnetic showers in the crystals but not in the hadronic calorimeters. In contrast to electrons, photons leave no trace in the tracker. Hadrons such as pions or neutrons produce large showers in the calorimeters. Unless they are very energetic, they dissipate all their energy before reaching the muon chambers.

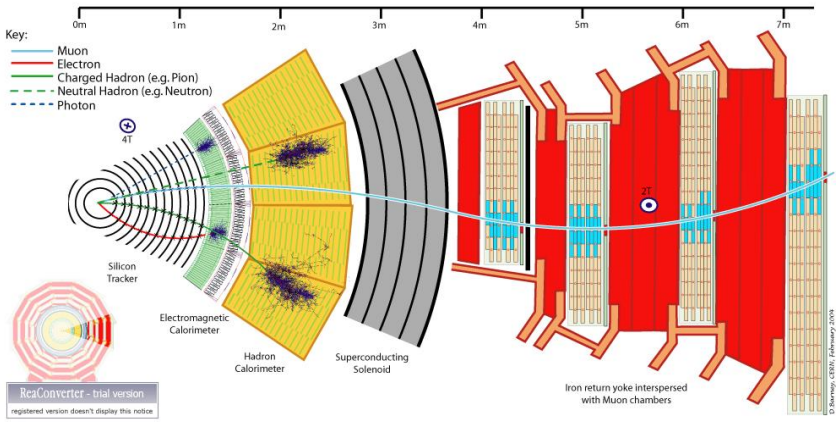


Figure 2.2: *Transverse CMS section showing the different detector layers.*

2.2 The Inner Tracking System

The inner tracking system of CMS [21] is designed to provide a precise and efficient measurement of the trajectories of charged particles as well as a precise reconstruction of secondary vertices. It surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. The CMS solenoid provides a homogeneous magnetic field of 3.8 T over the full volume of the tracker. At the LHC design luminosity of 10^{34} cm^{-2} s^{-1} there will be on average about 1000 particles from more than 20 overlapping $p-p$ interactions traversing the tracker for each bunch crossing, every

25 ns. Therefore a detector technology featuring high granularity and fast response is required. However, these features imply a high power density of the on-detector electronics, which in turn requires efficient cooling. There is also the requirement to minimize the amount of material in order to limit multiple scattering, bremsstrahlung, photon conversion and nuclear interactions. The CMS tracker is composed of a pixel detector with three barrel layers at radii between 4.4 cm and 10.2 cm and a silicon strip tracker with ten barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by endcaps which consist of 2 disks in the pixel detector and 3 plus 9 disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudorapidity of $\eta < 2.5$. With about 200 m^2 of active silicon area the CMS tracker is the largest silicon tracker ever built. The CMS tracker consists of 1440 pixel and 15148 strip detector modules. The performance of the tracker is illustrated in Fig.2.3 which shows the transverse momentum and impact parameter resolutions in the $r-\phi$ and $r-z$ planes for single muons with p_T of 1, 10 and 100 GeV/c, as a function of the pseudorapidity. The resolution on the p_T measurements remains below 2% up to $|\eta| < 1.6$.

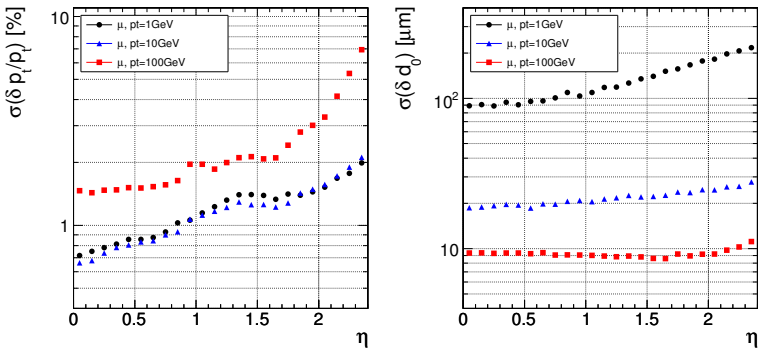


Figure 2.3: *Tracker tracks resolution on the transverse momentum (left) and the transverse impact parameter (right) as a function of pseudorapidity, for single muons with transverse momentum of 1, 10 and 100 GeV/c [24].*

2.2.1 The Pixel Detector

The detector consists of 4.4 millions of n-type silicon pixels with a size of 100 μm $\times$ 150 μm on a n-type silicon substrate. The sensor are read-out analogically and a spatial resolution of $\sim$ 10 μm for the $r - \phi$ coordinate and of $\sim$ 20 μm for the $r - z$ measurement is achieved interpolating the charge induced in nearby pixels, helped by the large Lorentz drift angle in the magnetic field. The charge sharing is enhanced in the endcaps, where the electric and magnetic fields are parallel, by tilting the detectors of 20°. The sensors are organised in three barrel layers at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, extending for a total length of 53 cm. In the endcaps two disks are placed at $|z| = 34.5$ cm and 46.5 cm extending in radius from 6 to 15 cm. The layout of the pixel detector is shown in Fig.2.4.

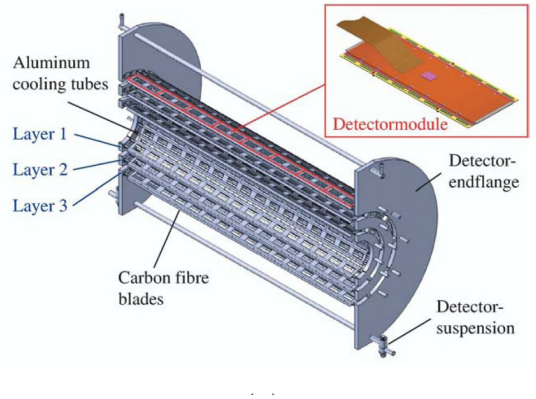


Figure 2.4: *Mechanical structure showing the three layers of the CMS pixel detector.*

The pixel detector is designed to provide at least two hits for tracks originating within $2\sigma_z$[1] from the nominal interaction point up to about

[1]Longitudinal size of the luminosity region of the LHC beam: $\sigma_z \sim 5$ cm.

$|\eta| < 2.2$. Moreover, stand-alone track reconstruction, which requires three hits per track, is also possible with good efficiency.

The three-dimensional resolution of the pixels will play a key role in the identification of b and τ-jets allowing the accurate measurement of the impact parameter and the reconstruction of primary and secondary vertices.

The short distance from the interaction point imposes special requirements on radiation hardness and will probably require the substitution of the pixel detector during the lifetime of the experiment.

2.2.2 The silicon Microstrip Detector

In addition to the pixel detectors, the tracker is composed of 10 layers of silicon microstrip detectors, whose layout is shown in Fig.2.5.

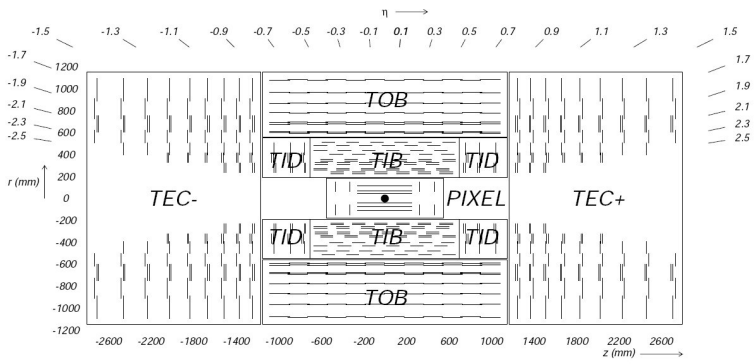


Figure 2.5: *Schematic cross-section of the Inner Tracking system showing also the η ranges of the different cross sections.*

The barrel region is divided into two parts: the TIB (Tracker Inner Barrel) and the TOB (Tracker Outer Barrel). The TIB is composed by four layers of n-type silicon sensors with a thickness of 320 μm and strip pitches varying from 80 to 120 μm. The first two layers are made with double-sided ("stereo") modules, composed by two detectors mounted back-to-back with

the strips tilted by 100 mrad. This kind of sensors provides a measurement in both $r-\phi$ and $r-z$ coordinates with a single point resolution between 23-34 μm and 230 μm respectively. The TOB is made of six layers. In this region the radiation levels are smaller and thicker silicon sensors (500μm) can be used to maintain a good S/N ratio for longer strip length and wider pitch. The strip pitch varies from 120 to 180 μm. Also the first two layers of the TOB provide a "stereo" measurement with a singlepoint resolution which varies from 35 to 52 μm in the $r-\phi$ direction and 530 μm in z.

The endcaps are divided into the TID (Tracker Inner Disks) and TEC (Tracker End Cap). The three disks of the TID fill the gap between the TIB and the TEC while the TEC comprises nine disks that extend into the region 120 cm $< |z| <$ 280 cm. Both parts are composed of wedge shaped modules arranged in rings, centred on the beam line, and have strips that point towards the beam line.

The entire silicon strip detector consists of about 15400 modules, which are mounted on carbon-fibre structures and housed inside a temperature controlled outer support tube. To enhance the radiation hardness of the system, the full detector will operate at a temperature of about $-20°C$.

2.3 The Calorimeters

Calorimeters play an important role in CMS. They measure energies of both neutral and electrically charged particles. Electrons, positrons and photons dissipate their energy by showering in the innermost section of the calorimeter, the electromagnetic calorimeter (ECAL). Motivated by the search for the two-photon decay of a Standard Model Higgs boson with a mass just above 114.4 GeV/c^2 (the limit reached at the LEP collider) a high-precision calorimeter was required. A homogeneous calorimeter made of lead tungstate crystals was chosen, which permits to achieve a mass resolution of better than 1% for low-mass Higgs. In order to obtain also a good rejection of $\pi^{0'}s$ decaying into two photons, a preshower detector is provided in the endcap regions. Hadrons are present in every collision. For the study of Supersymmetry, QCD and other topics a hadron calorimeter

is essential. It has been designed not only to measure energies of jets but also to provide good containment and hermeticity for the determination of missing energy. In order to reach a large rapidity coverage the barrel part (HB) and the endcap parts (HE) are complemented by two forward hadron calorimeters (HF) on each side of the detector. HB and HE are made of alternating brass-scintillator plates, and HF is made of steel absorber plates with embedded quartz fibres. An outer hadron calorimeter (HO), made of an additional layer of scintillator, has been included in the barrel region, just outside the coil, to maximise the number of interaction lengths[2] and thus to prevent punchthrough of hadronic showers into the muon system as much as possible.

2.3.1 The electromagnetic calorimeter

The ECAL[22] is a high-resolution, high-granularity detector made of lead tungstate ($PbW0_4$) crystals. Lead tungstate is a fast scintillator providing a small Molière radius and short radiation length.

The geometrical crystal coverage extends up to $|\eta| = 3.0$, as shown in Fig.2.6, but the high radiation dose of pile-up energy limits precision energy measurements to the range of $|\eta| < 2.6$.

The front face area of 22×22 mm^2 matches the Molière radius of lead tungstate (21.9 mm). This area corresponds to a granularity of $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ in the barrel region and 0.0175×0.0175 up to 0.05×0.05 in the endcaps. The crystals in the barrel section are 23 cm thick, which corresponds to 26 radiation lengths.
In the endcap region, 22 cm thick crystals are complemented by preshower detectors, that consist of lead/silicon detector layers and cover a range of $1.65 < |\eta| < 2.6$. These allow the identification of neutral pions and improve the position determination of electrons and photons.
In total, the electromagnetic calorimeter consists of 82728 crystals, 61200 of them in the barrel part, summing up to a volume of 11.18 m^3.

[2]The interaction length is the mean free path of a particle before undergoing an inelastic interaction.

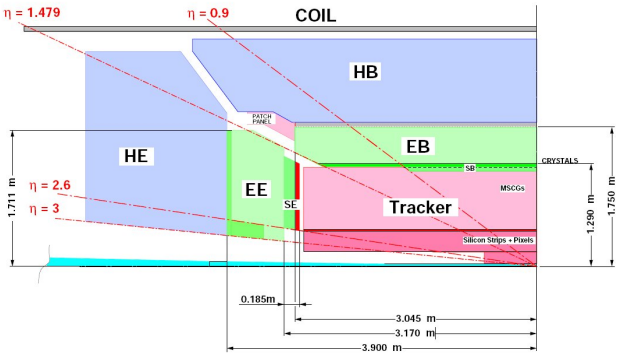


Figure 2.6: *Longitudinal view of one quater of the ECAL.*

The scintillation signals are detected by photodetectors. In the calorimeter barrel, avalanche photodiodes, which provide gain even within high transverse magnetic fields, are used. Vacuum phototriodes have been chosen for the endcaps, since they are able to resist the higher integrated radiation dose in these regions. A 16 bit readout chain records the photodetector signals. The corresponding energy range extends from about 30 MeV in the barrel part and about 150 MeV in the endcaps up to about 2 TeV per crystal.

A parametrisation of the energy resolution can be given by

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{b}{E}\right)^2 + c^2 \tag{2.1}$$

Here, a is a stochastic term including fluctuations in the shower containment and in the number of photo-electrons, b describes noise caused by electronics or pileup energy, and c is a constant characterising energy leakage and intercalibration errors. The values of the three constants measured on test beams are reported in Table 2.1. The different contributions as a function of the energy are shown in Fig.2.7.

Contribution	Barrel ($\eta = 0$)	Endcap ($\eta = 2$)
Stochastic term	2.7%	5, 7%
Constant term	0.55%	0.55%
Noise (low luminosity)	0.155 GeV	0.155 GeV
Noise (high luminosity)	0.210 GeV	0.245 GeV

Table 2.1: *Contributions to the energy resolution of ECAL.*

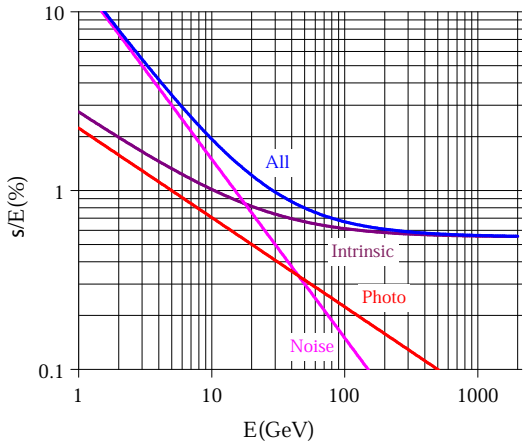


Figure 2.7: *Different contributions to the energy resolution of the ECAL. The curve labelled "intrinsic" includes the shower containment and a constant term of* 0.55%.

2.3.2 The Hadron Calorimeter

The goal of the hadron calorimeter (HCAL) [23] is to measure the direction and energy of jets, the total transverse energy and the balance of the transverse energy. High hermeticity is required for this purpose. For this reason, HCAL is arranged in a central calorimeter, which covers the range of $|\eta| < 3$, containing the Hadronic Barrel (HB), the Hadronic Endcaps

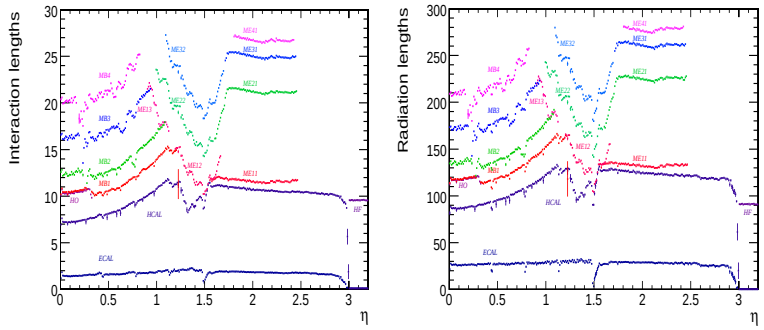
(HE), and the Outer Hadronic calorimeter (HO), complemented by the forward and backward calorimeters (HF) in the range of $3 < |\eta| < 5$. HCAL is a hadronic sampling calorimeter, consisting of brass and stainless steel absorbers and plastic scintillators. Its dynamic energy range of 5 MeV to 3 TeV allows the observation of single muons as well as the measurement of the highest possible particle energies.

The HB is divided into two half barrels, each containing 18 identical wedges. The wedges are made of absorber plates, complemented by 17 layers of plastic scintillators. The innermost and outermost absorber plates consist of stainless steel for stability reasons, the others are made of brass to maximise the hadronic interaction length. Due to its location inside the magnet coil, the thickness of the HB is restricted to 100 cm. The granularity of the scintillators in the barrel part is $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$.
Each of the Hadronic Endcaps consists of absorber plates and scintillators arranged in 10 sectors. There are 18 absorber layers of 80 mm thickness, each. As in the barrel part, the innermost and outermost layer are made of stainless steel, while the others consist of brass. The scintillator granularity is the same as in the barrel part, except for the highest η-regions. Here, the granularity matches that of the electromagnetic calorimeter.

The HO is placed outside the solenoid, enveloping the first iron absorber layer of the muon system. It contains one sampling layer in the endcap region, two layers in the barrel region and an additional layer in the range of $\eta < 0.4$. The Outer Hadronic Calorimeter is essential for the full containment of hadron showers. The total absorber thickness of the hadronic calorimeter corresponds in the barrel to 5.15 hadronic interaction lengths λ at $\eta = 0$ up to 9.1λ at $\eta = 1.3$ and averages at 10.5λ in the endcap regions (see Fig.2.8).

The HF calorimeters, positioned at a distance of about 11 m from the interaction point, are needed for identification and reconstruction of very forward jets. Due to the high radiation field in this region, quarz fibres were chosen as active material. These fibres emit Cherenkov light which is detected by photodiodes. They are placed between 5 mm thick steel absorber plates with a total thickness of 10 hadronic interaction lengths λ.

During test beam studies, the energy resolution of the CMS calorime-



(a) (b)

Figure 2.8: *Interaction (a) and radiation (b) length as a function of η.*

ter system (including the electromagnetic calorimeter) for hadrons in the energy range 30 GeV< E <1 TeV was found to be

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{100\%}{\sqrt{E}}\right)^2 + (4.5\%)^2 \tag{2.2}$$

2.4 The Superconducting Magnet

The 12000 tonnes heavy CMS Magnet [20] is the central device around which the experiment is built. It is a superconducting solenoid made of coils of wire that produce a uniform magnetic field of 3.8 T when current flows through them at a temperature of -268.5 °C.

CMS was designed to contain the strongest magnet possible with the idea that a higher strength field bends paths more and, combined with high-precision position measurements in the tracker and muon detectors, this allows accurate measurement of the momenta of even high-energy particles. The required accuracy is of about 1% for particles with $p = 100$ GeV/c. Moreover the unambiguous charge-assignment for muons with a momentum of ~ 1 TeV/c requires a momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1$ TeV/c.

The parameters of the solenoid are given in Table.2.2, while in Fig.2.9 the magnetic field layout is shown.

Parameter	Value
Field	3.8 T
Field in the yoke	2T
Inner Bore	5.9 m
Length	12.9 m
Number of Turns	2168
Current	19.5 kA
Stored energy	2.7 GJ
Hoop stress	64 atm

Table 2.2: *Parameters of the CMS superconducting solenoid.*

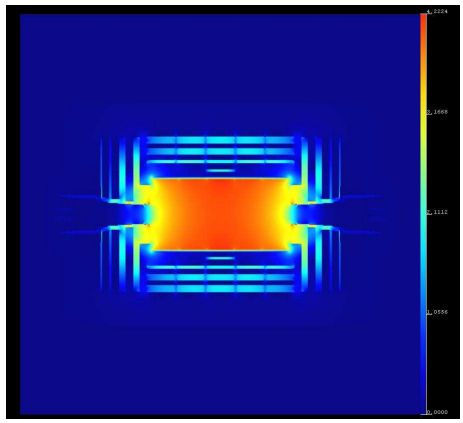


Figure 2.9: *Magnet field layout.*

2.5 The Muon System

Muons provide a very clear signal for expected new physics in a p_T range which spans three orders of magnitude (GeV to TeV). The CMS Muon System [25] aims to identify and measure their momenta very precisely. The muon trigger has to be very reliable in bunch crossing identification and has to be multifunctional for setting muon p_T thresholds. Muons tend to penetrate the calorimeters due to their masses 200 times higher than those of electrons, and their bremsstrahlung losses are negligibly small. Because of this fact the muon system is embedded in the outermost region of CMS in the return iron yoke of the magnet. The track sagitta due to the bending in the magnetic field, taking the point of interaction into account, is used to measure p_T. Four stations are placed in concentric cylinders around the beam line in the barrel region, and three disks are arranged perpendicular to the beam line in the endcaps (Fig.2.1).

The muon system consists of three independent subsystems. In the barrel ($|\eta| < 1.2$), where the track occupancy and the residual magnetic field are low, drift tube detectors (DT) are installed. In the endcaps ($0.8 < |\eta| < 2.4$), cathode strip chambers (CSC) are located, as detectors in this region have to cope with high particle rates and a large residual magnetic field between the plates of the yoke. Redundancy is obtained with a system of resistive plate chambers (RPC), that are installed in both the barrel and the endcaps. RPCs have limited spatial resolution, but fast response and excellent time resolution, providing unambiguous bunch crossing identification. They are also used to complement DTs and CSCs in the measurement of the p_T. The RPC system covers the region $|\eta| < 2.1$. Fig.2.10 shows the $|\eta|$ coverage of the muon system and the three muon chamber technologies.

As the high field is fundamental for the momentum resolution of the spectrometer, it also sets the environment in which the detector operates. The innermost endcap CSCs, the ME1/1 chambers, are exposed to the full field which, in this region, is almost entirely axial and uniform (see Fig.2.9). In the following CSC stations the field is no longer axial and uniform, however, the small drift space allows these detectors to limit the degradation of the chamber resolution. In the barrel region most of the flux

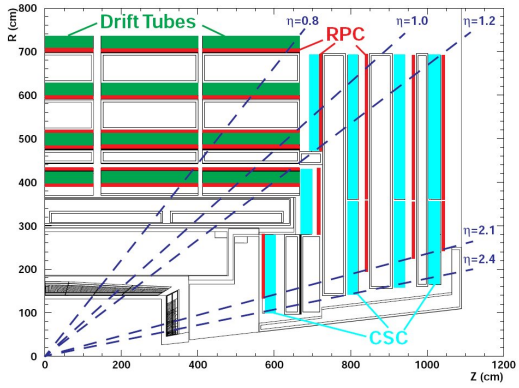


Figure 2.10: *Longitudinal cut of the CMS Muon System.*

is contained within the iron plates of the yoke where the axial component of the field reaches $\sim$ 1.8 T. The space where the DT chambers are placed should ideally be field-free. However in the iron gaps and at the end of the coil the residual magnetic field is far from being negligible. The axial and the radial components are shown in Fig.2.11 for the various barrel DT stations. There are spatially limited regions where the magnetic field in the radial direction reaches 0.8 T.

2.5.1 The Drift Tube Chambers

Muon detectors in the barrel do not operate in particularly demanding conditions, since the occupancy in this region is low and the magnetic field is well contained in the iron plates of the return yoke. For this reason, drift tubes were chosen.

The chamber segmentation follows that of the iron yoke, consisting in five wheels along the z axis, each one divided into 12 azimuthal sectors. The wheels are numbered from -2 to +2, sorted according to global CMS z axis, with wheel 0 situated in the central region around $\eta = 0$.

Within all wheels, chambers are arranged in four *stations* at different

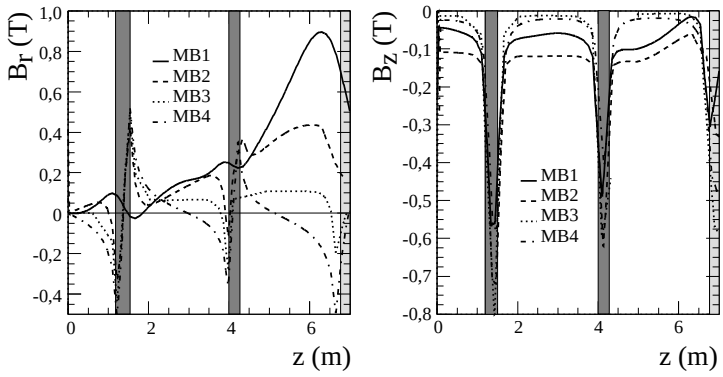


Figure 2.11: *Magnetic field components radial (B_r) and axial (B_z) with respect to the beam line in the different barrel stations as a function of the distance from the CMS centre in the axial direction (along the z-axis). The shaded areas are the gaps between the detector rings.*

radii, named MB1, MB2, MB3 and MB4 as shown in Fig.2.12 The first and the fourth station are mounted on the inner and outer face of the yoke; the remaining two are installed in slots within the iron plates. Each station consists of 12 chambers, one per sector, except for MB4 where 14 chambers are present. Sector numbering increases anti-clockwise when looking the detector from the positive z axis, starting from sector number 1 which contains the vertical chambers matching the positive x axis in the CMS global reference frame. The additional two chambers in MB4 are numbered 13 and 14 according to the scheme (see Fig.2.12)

The basic detector element is a drift tube cell, whose section is shown in Fig.2.13. Cells have a size of 42 $\times$ 13 mm^2. A layer of cells is obtained by two parallel aluminium planes and by "I" shaped aluminium beams which define the boundary of the cells and serve as cathodes. I-beams are insulated from the planes by a 0.5 mm thick plastic profile. The anode is a 50 μm stainless steel wire placed in the centre of the cell. The distance of the track from the wire is measured by the drift time of electrons produced by ionisation. To improve the time-distance linearity, additional field shaping

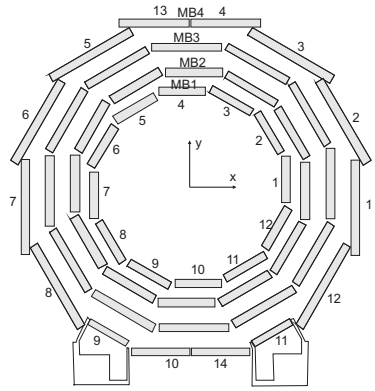


Figure 2.12: *Numbering of stations and sectors.*

is obtained with two positively biased insulated strips, glued on the planes in correspondence to the wire. Typical voltages are +3600 V, +1800 V and -1200 V for the wires, the strips and the cathodes, respectively. The gas is a 85%/15% mixture of Ar/CO_2, which provides good quenching properties and a saturated drift velocity, of about 5.4 cm/μs. The maximum drift time is therefore $\sim$ 390 ns, *i.e.* 15 bunch crossings. A single cell has an efficiency of about 99.8% and a resolution of $\sim$ 180 μm.

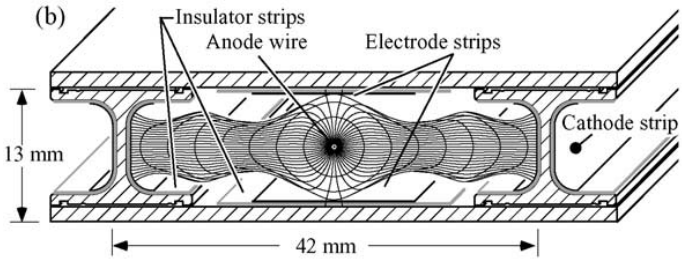


Figure 2.13: *Section of a drift tube cell.*

Four staggered layers of parallel cells form a *superlayer*, which provides the measurement of a two-dimensional segment solving the left-right ambiguity of a single layer by means of pattern recognition. Also, the information of a superlayer can be used by the trigger to identify the bunch crossing originating a segment with no need of external input, using a generalisation of the meantimer technique [26].

A chamber is composed by two superlayers measuring the $r-\phi$ coordinate, with the wires parallel to the beam line, and an orthogonal superlayer measuring the $r-z$ coordinates. The latter is not present in the outermost station (MB4). The wire length varies, depending on the superlayer type and on the station, from about 2 m to 3 m. The superlayers are glued on the two sides of an honeycomb spacer which ensures a $\sim$ 15 cm lever arm as well as the required stiffness to the structure. A cross-sectional view of a chamber is shown in Fig.2.14.

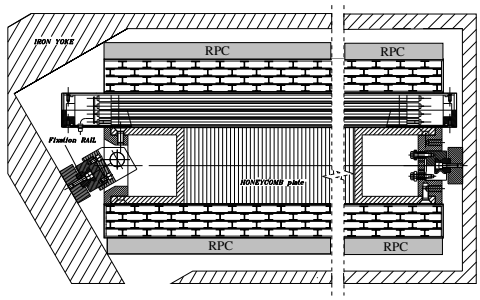


Figure 2.14: *Schematic view of a MB1/MB2 DT chamber. The two resistive plate chambers of the muon station are also shown.*

The convention on chamber local reference frames, used in the CMS reconstruction software, is shown in Fig.2.15. Superlayers are numbered from 1 (innermost $r-\phi$) to 3 (outermost $r-\phi$). Number 2 is thus always the $r-z$ superlayer, when present. Layers are numbered from 1 (innermost) to 4 (outermost). A local reference frame is associated to each layer; the origin is placed at the centre of the layer, the x axis along the measured

coordinate and the y axis along the wires, pointing towards the superlayer's front-end side. Cell number increases with x increasing. Superlayers have a local reference frame oriented in the same way as the layers, but with the origin at the superlayer centre. The chamber reference frame has the origin in the chamber centre and is oriented as the $r-\phi$ superlayer reference frames, *i.e.* with the x axis along the coordinate measured by the $r-\phi$ superlayers and the y axis along the coordinate measured by the $r-z$ superlayer, pointing towards the $r-\phi$ front-end side. With this choice of reference frames, the z axis of all local reference frames points toward the interaction region.

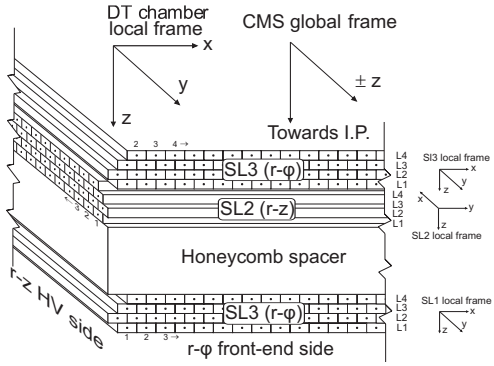


Figure 2.15: *Scheme of a DT chamber, showing the conventions on superlayers, layer and wire numbering and the orientation of reference frames.*

2.5.2 The Cathode Strip Chambers

Cathode Strip Chambers are used in the endcap regions where the magnetic field is very intense and inhomogeneous and where the charged particle rate is high. CSCs are multiwire proportional chambers defined by two cathode planes, one segmented into strips, and an array of anode wires running across, in between (Fig.2.16). An avalanche developed on a wire induces

a charge on several strips of the cathode plane, and interpolation between adjacent strips gives a very fine spatial resolution of 50 μm, which is used to measure the ϕ-coordinate. Simultaneously, the wire signals are read out, directly, and used to measure the radial coordinate with a coarse precision of approximately 0.5 cm. The closely spaced wires make the CSC a fast detector suitable for triggering. The basic module of the CSC system is a chamber consisting of six layers in order to provide high efficiency and a robust pattern recognition that can reject non-muon background. Combining multiple layers also improves the timing resolution so that the correct bunch crossing can be assigned with over 99% efficiency. The cathode planes are formed by honeycomb panels with copper clad skins, while the 9.5 mm thick gas gaps are filled with a mixture of 30% Ar, 50% CO_2 and 20% CF_4. Chambers are of trapezoidal shape with strips running radially (strips have constant width). They cover sectors of 10° or 20° and have a maximum dimension of 3.5 m to 1.5 m. The chambers are placed in four disks (stations ME1 to ME4) perpendicular to the beam axis. The inner rings of stations ME2 to ME4 consist of 18 chambers while all the other rings consist of 36 chambers. The chambers overlap in order to avoid dead areas, except in the outermost ring of the first station (ME1/3) which has gaps between the chambers. The innermost chambers of the first station (ME1/1) are slightly different from the other chambers as they operate in the highest magnetic field (up to 3 T) and under the highest radiation dose.

2.5.3 The Resistive Plate Chambers

A Resistive Plate Chamber is capable of tagging the time of an ionizing event in a much shorter time than the 25 ns between two consecutive LHC bunch crossings. Therefore, a fast dedicated muon trigger device based on RPCs can identify unambiguously the relevant bunch crossing to which a muon track is associated even in the presence of the high rate and background expected at the LHC. RPCs are parallel-plate detectors filled with gas having an excellent time resolution of approximately 1 ns, comparable to that of scintillators. A RPC consists of two parallel plates made of very high resistivity plastic material (bakelite), separated by a gas gap of

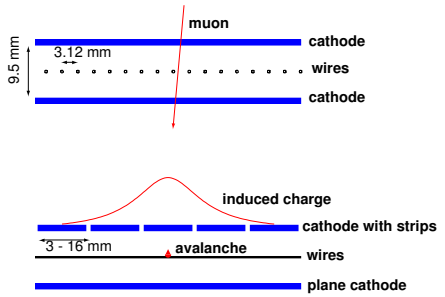


Figure 2.16: *Orthogonal section of one CSC chamber.*

a few millimeters (Fig.2.17). The plates are coated with graphite on the outside which forms the high voltage electrodes. The read-out is performed by means of aluminum strips separated from the graphite coating by an insulating PET (polyethylene terephthalate) film. The RPCs are operated with a gas mixture of 95% $C_2H_2F_4$ and 5% i-C_4H_{10}.
In the barrel six layers of RPCs are used. They are attached to either side of the DT chambers in the inner two stations and only to one side in the outer two stations. This allows to have a minimum of four measurements also for low-p_T muons which do not reach the outer two stations. The strips are oriented parallel to the beam-line with each strip covering 5/16 degrees in ϕ. In the endcaps, four layers of RPCs are attached to the four CSC disks, covering a range up to $|\eta| < 2.1$. The strips are trapezoidal in shape and oriented radially: they also cover 5/16 degrees in ϕ per strip (Fig.2.10).

2.6 Forward Detectors

2.6.1 CASTOR

The CASTOR [8] (Centauro And Strange Object Research) detector is a calorimeter made of quartz-tungsten in the very forward rapidity region before the HF layer. It is developed in the baryon-free mid-rapidity region

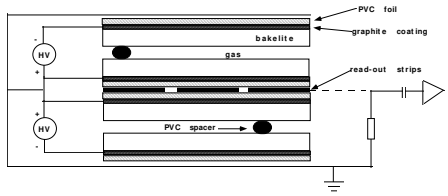


Figure 2.17: *Scheme of a double-gap Resistive Plate Chamber.*

for the heavy ion physics program, and also for the diffractive and low-x physics program in $p-p$ collisions. CASTOR has been installed at 14.38 m from the interaction point placed in the $5.2 < |\eta| < 6.6$ range. The calorimeter is constructed in two halves surrounding the closed beam pipe. The typical total and electromagnetic energies in the CASTOR acceptance range (about 180 TeV and 50 eV, respectively, for simulated Pb-Pb collsions at 5.5 TeV) can be measured with a resolution better than 1%. The main advantages of quartz calorimeters are radiation hardness, fast response and compact detector dimensions.

2.6.2 Zero Degree Calorimeter (ZDC)

For diffractive studies two zero degree calorimeters [8] with pseudorapidity coverage of $|\eta| \geq 8.3$ for neutral particles are placed in the CMS forward region. Each ZDC has two independent parts: the electromagnetic (EM) and hadronic (HAD) sections. Two identical ZDCs will be located between the two LHC beam pipes at 140 m on each side of the CMS interaction point. During heavy ion running the combined (EM + HAD) calorimeter should allow the reconstruction of the energy of 2.75 TeV spectator neutrons with a resolution of 10-15%. Sampling calorimeters using tungsten and quartz fibers have been chosen for the detection of the energy in the ZDCs with a design similar to HF and CASTOR.

Chapter 3

The Drift Tube calibration procedure

As already discussed in the previous chapter, the main goal of the CMS spectrometer is to identify and measure with great precision the momenta of muons. In particular the DT system is designed to reconstruct muon tracks with a correct charge assignment up to the TeV region. Moreover Drift Tubes are used to provide an independent first level trigger selection, with a fast muon identification and accurate transverse momentum measurement, besides excellent bunch-crossing assignment.

A fundamental ingredient in the DT system is the calibration of the signal arrival time and drift velocity used as input to the local hit reconstruction within the cells. Therefore it influences directly the precision of the tracking resolution and the track efficiency. After a brief summary of the main DT detection principles (Sez.3.1), the parameters computed during the calibration process are described in Section 3.2 and 3.3. Finally, in Section 3.4 is presented the DT calibration work-flow, performed within the CMS computing framework, including the monitoring of the conditions.

All the results presented in this chapter have been obtained with data collected during the high statistics cosmic ray data taking period, the so-called Cosmic Run At Four Tesla (CRAFT) [27]. During CRAFT, most of

the CMS sub-detectors participated, and data were taken with the magnetic field switched both on and off. About 300 million events were collected with a magnetic field inside the solenoid of 3.8 Tesla. The DT system was the primary trigger source for most of the collected events. All the sectors and wheels of the system participated and after proper chamber synchronization between upper and lower sectors, the DT trigger provided a stable cosmic muon rate of about 240 Hz during the entire data taking period.

3.1 The DT detection principles

The CMS drift tubes detect charged particles by recording the ionization electrons produced while they cross the chamber. The drift time of such electrons in a properly shaped electrostatic field is measured to get the information about the spatial coordinates of the ionizing event.

As explained in Section 2.5.1 the electric field is in fact shaped in order to have a linear space-time relationship over the entire volume, except for the region close to the central anode wire, where it increases as $1/r$. Near the anode the electrons are accelerated enough to produce secondary ionization and hence an avalanche. A gain of about 10^5 is achieved. The CO_2 present in the gas admixture acts as quencher. Thanks to its large photoabsorption cross section, it absorbs photons produced during the avalanche development, keeping the avalanche region limited. The anode signal, induced by the electron avalanche, is amplified and discriminated against a configurable threshold by the front-end electronics and it is then sent to the read-out electronics performing the TDC measurement.

Electrons produced at a time t_{ped} by the incoming particle migrate toward the anode with a velocity v_{drift} and reach the anode at a time t_{TDC}, which is the time measured by the TDC. The distance of the track from the anode wire is therefore given by:

$$x = \int_{t_{ped}}^{t_{TDC}} v_{drift} \cdot dt \tag{3.1}$$

The time pedestals are extracted directly by the measurement distributions

with the calibration procedure described in Section 3.2.1.

The computation of the correct drift velocity is instead more complex because different kind of cell non-linarities have to be taken into account. The major factors which can influence the drift velocity are:

- electric field strength and direction and gas pressure. To limit the dependence from these two factors, the working point is chosen to have a value of E/p corresponding to a saturated drift velocity of 1.9 kV/cm.

- gas composition and temperature. The pressure, temperature and purity of the mixture are continuously monitored. In particular the oxygen concentration is always kept below 500 ppm;

- presence of external factors such magnetic stray fields, which modify the drift properties of an electron swarm. The Lorentz force applied to each of the moving charges bends its trajectory. Therefore, the resulting field lines are distorted, as shown in Fig.3.1. This effect is present in the whole DT system, because the residual magnetic field in the cell volume is not homogeneous, increasing in the inner stations where the magnetic field is higher.

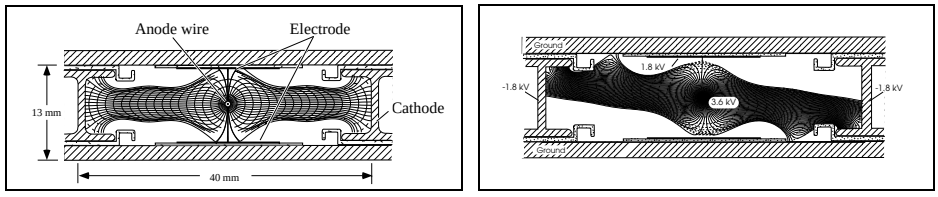


Figure 3.1: *Transverse view of a Drift Tube cell in the barrel. The drift lines and the isochrones for a typical voltage configuration of the electrodes and representative gas mixture are shown for the case of zero magnetic field (left) and a 0.45 T magnetic field parallel to the anode wires (right).*

Moreover for tracks inclined with respect to the normal to the layer in the measurement plane the time-distance relation is modified by the fact

that the electrons having the shortest drift time are not those produced in the mid plane of the drift cell. The effective drift-velocity is therefore higher. No deviation from linearity is observed for tracks inclined in a plane parallel to the wire.

Two algorithms are available in the CMS reconstruction code. The first reconstruction algorithm is based on the assumption of a constant drift velocity within the entire cell. In this case, Eq.3.1 becomes:

$$x = (t_{TDC} - t_{ped}) \cdot v_{drift}^{EFF} = t_{drift} \cdot v_{drift}^{EFF} \tag{3.2}$$

where v_{drift}^{EFF} is the effective average drift velocity.

The goal of the calibration procedure is in this case to determine the time pedestal (t_{ped}) which is needed to extract the drift time (t_{drift}) from the TDC measurement (t_{TDC}), and the average drift velocity (v_{drift}^{EFF}).
The value of v_{drift}^{EFF} is computed on limited spatial regions of CMS where the drift velocity conditions are assumed to be uniform, using the mean-time technique [26]. This method have been used for the DT calibration during the CRAFT period. Results are shown in the following.

The second reconstruction algorithm is based on a parameterization of the cell response [28] obtained with GARFIELD [29]. This parameterization includes the dependence on the track impact angle, α. and on the stray magnetic field $\mathbf{B}$:

$$x = f((t_{TDC} - t_{ped}), \alpha, \mathbf{B}) \tag{3.3}$$

In this case the only quantity to be calibrated is t_{ped}, as the dependency on the relevant parameters is already accounted for by the parameterization.

3.2 Calibration parameters

3.2.1 Calibration of the Time Pedestals

A DT measurement consists in a TDC time which contains several contributions other than the ionization electron drift time:

- propagation time of the signal along the anode wire;
- delays due to the cable length and read-out electronics;
- delays due to the first level trigger latency;
- time-of-flight of the muon from the interaction point to the cell.

The goal of the calibration of the time pedestals is to determine the average value of these contributions needed to estimate the drift-time. This delay is usually referred as t_{trig} because of its dependence on the trigger latency.

The extraction of the drift time from the TDC measurement is performed in several consecutive steps.

- *Inter-channel synchronization.*
  First, a dedicated hardware calibration procedure, described in Section 3.2.1.1, allows to compute the relative offset among different wires, called t_0, due to different cable length and delays in the read-out electronics circuits.

- *Absolute pedestal determination*
  Once the channels are synchronized, it is possible to compute the absolute offset of the drift time distribution. This absolute pedestal is accounted by the t_{trig} which is estimated from the distribution of the TDC measurements superlayer by superlayer. The calibration of the t_{trig} is described in Section 3.2.1.2.

Note that the determination of these two delays does not solve completely the problem of the synchronization of the TDC measurements: the t_{trig} offset only accounts for the average time-of-flight and for the average delay for the signal propagation along the wire. Further corrections for these quantities are applied on an event by event basis during the reconstruction procedure, after the three-dimensional hit position within the chamber is known and the hits are associated into 3D track segments.

Specifically, if the t_{trig} is computed for a full superlayer uniformly illuminated:

- the 3D position obtained from the segment extrapolation to the hit plane, if available, is used to correct the TOF with respect to the superlayer center;
- the hit coordinate along the wire is used to correct the propagation time with respect to the middle of the wire, assuming a propagation velocity of 0.244 m/ns, as directly measured on test beam data [30].

These corrections can be as high as $\sim$ 2 ns for the TOF and $\sim$ 6 ns for the signal propagation delay and they can be adapted or switched off in case of different running conditions.

3.2.1.1 The inter-channel synchronization

The synchronization among different channels is calculated for each chamber, in order to correct for the different signal path lengths of trigger and readout electronics. This offset is computed for each channel and it is a fixed number since it depends only on cable/fiber lenghts.

The inter-channel synchronization is performed through special test-pulse calibration runs: a test-pulse is generated at the input of the front-end chip for 4 wires (one per layer) at once, simulating a muon crossing a Superlayer. The test-pulse signal is injected simultaneously every 16 channels in the entire DT system to allow a fast scan of all the cells. Such dedicated calibration runs will be taken during the normal physics data taking exploiting the collision-free interval of the LHC beam structure, called "abort gap".

The resulting distribution of the measured times is extremely narrow, with a width of 1 or 2 TDC counts[1]. The average time value of each channel is stored in a database to be used later during the reconstruction. Note that the only relevant information contained in these delays is the relative difference among different channels since the absolute offset is just determined by the particular latency of the trigger used to read out the test-pulses.

[1]For the HPTDC chip used in the DT electronics 1 TDC count = 0.7812 ns.

Fig.3.2 shows an example of distribution of t_0 constants for representative layers of the three super-layers of a chamber as a function of the channel number. The t_0 synchronization correction is always below 10 ns (1 TDC count corresponds to 0.78 ns). The standard deviation is about 1 ns, for all channels. This is compatible with the precision of the electronic chain. The amount of delay observed in the figure corresponds approximately to the distance of the front-end boards present inside the chamber volume to the read-out boards. The stepping function shape, observed in Fig.3.2, more explicitly in the superlayer 1 and 3, shows the distance of groups of channels connected to the same front-end boards to the read-out board.

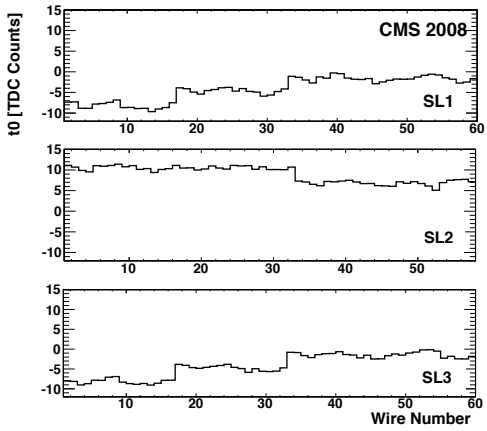


Figure 3.2: *Distribution of inter-channel synchronization constants calculated from a test-pulse run. Results are shown for three representative layers, belonging to each of the three super-layers of chamber MB3 in Sector 9.*

A final additional correction must be considered in the t_0 calibration. In each superlayer the pulse injection lines of odd and even layers are different. Therefore the two kinds of layers must be synchronized, on average, among them before computing the relative t_0 in each channel.

The inter-channel calibration performed subtracting the t_0 from pulses to the TDC measurements allows to compute the absolute offset of the drift-time distribution for a superlayer at a time.

3.2.1.2 The time pedestals computation

Since the TDC measurements of the different channels in a chamber have already been synchronized by subtracting the t_0 offset, the t_{trig} can be computed with every possible granularity within the chamber. The usual choice is to compute it superlayer by superlayer, as a compromise between accuracy in accounting for the average TOF and the quantity of available data.

Different contributions related to the hit position, track inclination, trigger latency, and cable lengths of the read-out electronic take part to the generation of the Time Pedestals. The contributions to the TDC values are included in the following expression:

$$t_{TDC} = t_0 + t_{TOF} + t_{prop} + t_{L1} + t_{drift} \tag{3.4}$$

Where:

- t_0 is the inter-channel synchronization to equalize the response of all the channels at the level of each chamber as described in Section 3.2.1.1;
- t_{TOF} is the Time-Of-Flight (TOF) of the muon from the interaction point to the cell;
- t_{prop} is the propagation time of the signal along the anode wire;
- t_{L1} is the time latency due to the first level trigger:
- t_{drift} is the drift time of the electrons from the ionization cluster within the cell.

The pedestal can be estimated directly from the distribution of the digi times, which is usually referred as the *time box*. The drift time varies

between 0, for muons passing close to the anode wire, to about 380 ns, for muons passing close to the cathode, but the full drift time distribution is shifted of a fixed value which is the t_{trig} pedestal. In case of a perfect linear cell behavior the distribution would have a regular box shape with a peak at low time values due to the non-linearity close to the anode wire. Some non linearity effects related to the electric field distribution inside the drift cell are nevertheless present in the response of the cells and are further enhanced due to the track inclination and the presence of residual magnetic field.

In order to compute the pedestal it is necessary to find a feature of this distribution which can be identified in an unambiguous and automatic way. Therefore a robust method [31], not sensitive to noise and spikes due to the read-out electronics, have been developed. This method is based on a fit of the rising edge of the drift time distribution (the so-called *error function*):

$$f(t) = \frac{1}{2} I \left[1 + \mathrm{erf}\left(\frac{t - \langle t \rangle}{\sigma\sqrt{2}} \right) \right] \tag{3.5}$$

where the normalization I, the standard deviation σ and the mean $\langle t \rangle$ are free parameters of the fit. In Fig.3.3 an example of this fit is shown with the continuos red line.

The inflection point of the rising edge of the time box does not directly represent the time pedestal of the distribution, but can be related to it by defining

$$t_{trig} = \langle t \rangle - k \cdot \sigma \tag{3.6}$$

where the mean of the integral of the Gaussian ($\langle t \rangle$) is located in the middle of the rising edge and the standard deviation (σ) is an indication of its slope. The optimal value of the t_{trig} should correspond to the beginning of the rising edge; this is achieved with Eq.3.6 by a properly computation of the k factor, as explained in the following.

The distribution of $\langle t \rangle$ and σ are shown in Fig.3.4 for a CRAFT run with the magnet field at 3.8 T. The values are computed for the most internal superlayer ($\phi 1$) as a function of chamber type and sector. The values of the $< t >$ parameter, which is the mean of the fit to the time box rising edge,

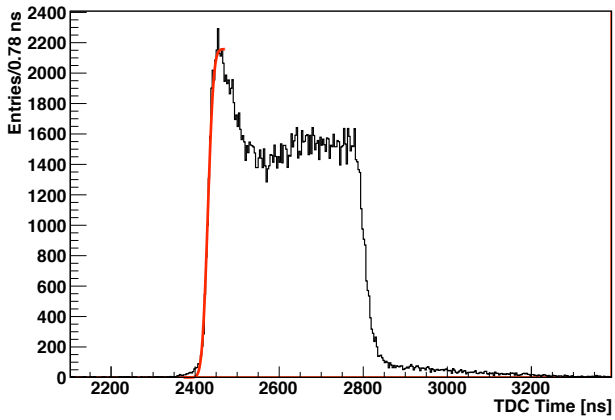


Figure 3.3: *Distribution of the signal arrival time recorded by the TDC; the arrival time in all the cells from a single superlayer in a chamber are superimposed, after the t_0 synchronization. The red line indicates the fit to the time box rising edge with the integral of a Gaussian.*

are approximately constant within each chamber type and sector number. A periodic structure is evident which corresponds to the TOF of the cosmic muon from the upper sector to the lower sector. In addition, the events triggered by the top sectors may be detected also by other non triggering sectors which will present a less precise time pedestal and consequently a worse determination of the calibration constants. The distribution of the slope, is an indication of the precision which the calibration procedure can reach. On average a value of 10 ns is observed for all superlayers in all wheels, apart for superlayer in vertical sectors, where the statistics is limited and the muon crossing angles are large.

Once $\langle t \rangle$ and σ are computed, the calibration process proceeds with the validation step studying the effect of these constants on the reconstruction algorithm. The hit position is determined assuming at this stage the t_{trig} constants computed in the first step, a constant drift velocity of 53.4 μm/ns and a constant k factor of 0.7. The analyzed quantities are the residuals computed as the difference between the distance of the hit and the 3D

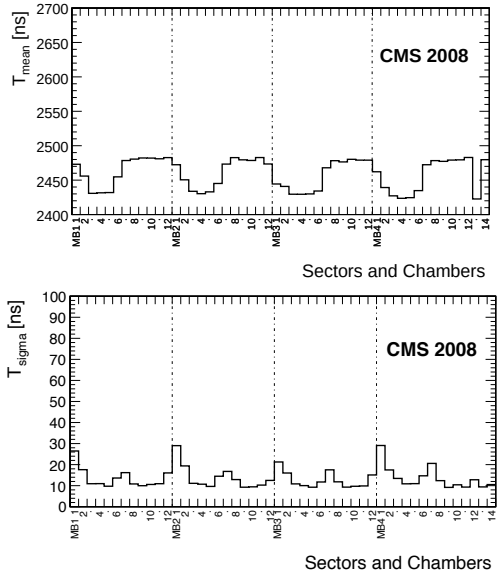


Figure 3.4: *Distribution of the mean (top) and standard deviation (bottom) of the fit to the rising edge of the time box using the integral of a Gaussian for the $r-\phi$ superlayers for a representative wheel. The triggering sectors (3,4,5 and 9,10,11) are synchronized among them.*

segment reconstructed fitting all the hits inside each chamber. These distributions show offsets which may be as large as 200 μm indicating that, at this level, the calibration of the space-time relationship has not yet reached the required precision.

A correction to the Time Pedestals is applied by changing the k factor, then those new offsets are divided by the constant drift velocity and finally the residuals are fitted again. With this correction the mean of the residual distribution, shown in Fig.3.5 (top), has an offset from zero compatible with the precision that can be reached with cosmic event structure.

The distribution of the standard deviations of the fit to the residuals is shown in Fig.3.5 (bottom). This represents the resolution obtained with

the local pattern recognition performed in the calibration procedure with values in the interval of 400-600 μm which is a factor 2 worse than the nominal resolution. This result is related to the uncertainty on the muon arrival time inside the 25 ns time window associated to the first level trigger. In fact the DT system was designed for the detection of muons coming from collision which are synchronized with the LHC clock while the arrival time of cosmic rays is uniformly distributed inside a clock cycle. This characteristic represents an intrinsic limitation for the calibration procedure and it affects the resolution obtained by the local pattern recognition. A more precise value of the drift velocity parameter together with the expected chamber track resolution, of about 200 μm, is reached in the offline local reconstruction where the muon arrival time and the drift velocity can be treated as free parameters in the refitting of each segment as described in Section 3.2.3.

3.2.2 Calibration of the Drift Velocity

As already discussed in Section 3.1 the drift velocity depends on many parameters such as the gas purity and conditions and the electrostatic configuration of the cell. Moreover, the bending induced by the magnetic field and the effect of the track angle can have an influence on the effective drift velocity modifying the drift path of the electrons. The main aim of the drift velocity calibration is to find the best effective drift velocity in each geometrical region of the DT system. A compromise between uniform experimental conditions and the use of geometrical regions which are as close as possible to the ones used for the t_{trig} calculation, has led to the choice of a superlayer granularity for the estimation of the drift velocity. The algorithm is based on the *mean timer* computation, described in Section 3.2.2.1, which permits to find the average velocity for different groups of cells. This technique allows to estimate the maximum drift time and therefore the average drift velocity in the cell. Moreover, it also allows to measure the cell resolution, which can be used as an estimate of the uncertainties associated to each measurement. This calibration procedure is discussed in Sect. 3.2.2.2.

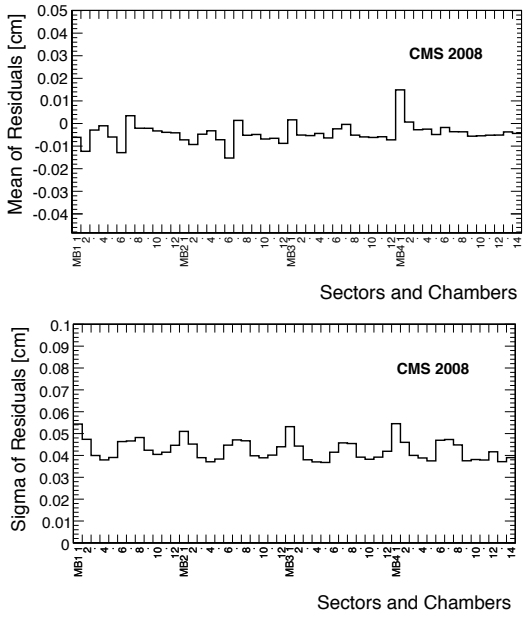


Figure 3.5: *Distribution of the mean (top) and standard deviation (bottom) of the Gaussian fit to the residuals between the reconstructed hits and the reconstructed local segments for the $r-\phi$ superlayers for a representative wheel after applying the correction from the offset of the residual distribution. Other wheels show similar results.*

3.2.2.1 The Mean Timer method

The meantimer formulas are relations among the drift times produced by a track in the different layers of a superlayer and the maximum drift time in the semi-cell.

The mathematical expression of the mean timer depends on the track angle and on the pattern of cells hit by the track. The most trivial relation is the one for a track crossing a semi-column of cells (see Fig.3.6): from simple geometric considerations it descends that the drift times in three

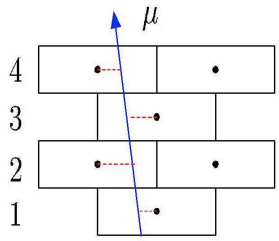


Figure 3.6: *Schematic view of a muon crossing a semi-column of cells.*

consecutive layers $(t_1.t_2, t_3)$ are related by:

$$T_{max}^{123} = \frac{t_1 + t_3}{2} + t_2 \tag{3.7}$$

where T_{max}^{123} represents the maximum drift time in the semi-cell. Note that this is true only if no delta-rays are produced: the signal generated by the secondary would mask the one produced by the muon resulting in a shorter drift-time and therefore in a smaller T_{max}^{123}.

The meantimer relations for different track angles and patterns of hit cells are listed in [31]. It should be noted that not all the track geometrical configurations can be used because in some cases the relation between drift times is independent of T_{max}. The proper meantimer formula is chosen track by track, using the direction and position information provided by the three-dimensional segments in a superlayer. This implies an iterative calibration procedure, starting with values of the drift velocity and of t_{trig} that already result in efficient pattern recognition and segment reconstruction.

3.2.2.2 The velocity drift computation

The calibration procedure consists of the following steps:

- a fit with a Gaussian of the meantimer distributions for each pattern j to estimate the mean value T_{max}^{j}, the sigma σ_T^j and the error on the

mean $\sigma_T^j/\sqrt{N_j}$ (where N_j is the number of entries in the distribution). The T_{max}^j contains the information about the average drift velocity in the different regions of the cell, since it is computed using drift times produced by hits in all the gas volume. An example of T_{max} distributions for different segment patterns is shown in Fig.3.7;

- the weighted average of the T_{max}^j is computed where the weights are taken as $N_j/(\sigma_T^j)^2$:

$$\langle T_{max} \rangle = \frac{\sum_j \frac{T_{max}^j}{(\sigma_T^j)^2} N_j}{\sum_j \frac{N_j}{(\sigma_T^j)^2}} \tag{3.8}$$

  This accounts for the relative importance of the different cell patterns in the computation of the maximum drift time.

- once the $\langle T_{max} \rangle$ is computed it is straightforward to find the average drift velocity through the relation:

$$v_{drift} = \frac{L/2}{\langle T_{max} \rangle} \tag{3.9}$$

  where L is the width of the cell. An example of drift velocity distribution computed with the superlayer granularity is shown in Fig.3.8 for each chamber/sector and for all the wheels. The figure shows that the drift velocity has approximately a constant value of 54.3 μm/ns, although with some fluctuations originating from the limitation of the calibration procedure applied to cosmic ray events. As already seen for the case of Time Pedestal, these fluctuations are originating mainly from the convolution of the non bunched structure of the arrival time of cosmic muons. The uncertainties on the time calibration influence directly the drift velocity determination. An error of 1 ns produces a 0.25% error in the drift velocity calculation, and the uncertainty observed in the t_{trig} calculation on CRAFT data (Fig.3.4) is of the order of 10 ns which corresponds to an error of 2.5% for the drift velocity. The fluctuations observed in Fig.3.8 show this uncertainty.

For this reason the residuals calculated with the v_{drift} and the t_{trig} constants properly calibrated, do not present significant improvement with respect to the ones shown in Fig.3.5.

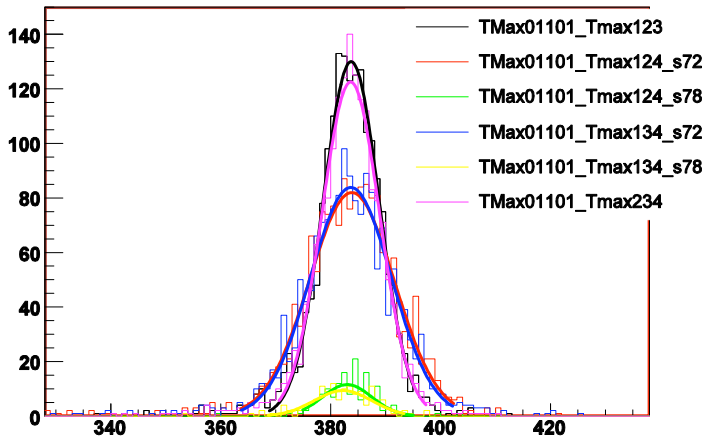


Figure 3.7: *An example of T_{max} distributions for different track crossing angles shown by different colors.*

Moreover this calibration procedure, through the mean timer technique, allows the estimation of the cell resolution and therefore of the uncertainties on the reconstructed distance.

The sigma of the mean timer distributions (σ_T^j) are a measurement of the resolution on the T_{max}^j. They can be therefore used to estimate the uncertainty on the measurement of the drift times (σ_t^j) with a relation which depends on the particular formula used to compute the mean timer. In case of tracks crossing a semi-column of cells, given the mean timer relation of Eq.3.7, the time resolution can be computed as:

$$\sigma_t^j = \sqrt{\frac{2}{3}} \cdot \sigma_T^j \tag{3.10}$$

which is valid under the assumption that the uncertainties are the same for the three layers used in the mean timer computation.

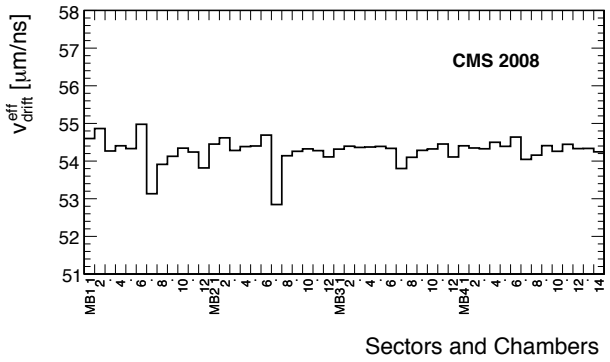


Figure 3.8: *Distribution of the drift velocity computed using the mean timer method for a run with B = 3.8 T. Results fo r the r − φ superlayers of each chamber/sector of a representative wheel. The other wheels show similar results.*

Since the cell resolution depends on the track angle, an average effective value is computed averaging the different values obtained for the contributing cell patterns weighted on the number of entries in each mean timer histograms:

$$\langle \sigma_t \rangle = \frac{\sum_j \sigma_t^j \cdot N_j}{\sum_j N_j} \tag{3.11}$$

The resolution on the reconstructed distance is therefore given by:

$$\sigma_d = v_{drift} \cdot \langle \sigma_t \rangle \tag{3.12}$$

This value is used during the reconstruction to assign the uncertainties to the one-dimensional hit in the gas volume. These uncertainties include the effect of the cell non-linearities (as those shown in Fig.3.1) only on average, therefore their dependence on the distance from the wire cannot be taken into account with this method.

3.2.3 Analysis on the drift velocity

The drift velocity obtained with the calibration procedure described in the previous section is derived from the measurements of the drift time and, as already mentioned, it is limited by the uncertainty on the arrival time of the cosmic ray events.

A more detailed analysis of the drift velocity [32] is presented here taking into account the precise 3D space-time relationship for the hit reconstruction. In particular, it considers the influence of the residual magnetic field as a function of the wire length in each cell.

Two different methods have been considered for measuring the drift velocity. The first uses the mean timer method for the drift velocity calibration procedure as has been described in Section 3.2.2.1. The second method relies on a muon track fit done in two steps. First the reconstruction software performs a pattern recognition algorithm and reconstructs tracks using the nominal drift velocity value (54.3 μm/ns). In the second step a new fit is performed using the associated hits to the track leaving as free parameters the drift velocity and the time of passage of the muon through the chamber. The method can be applied to the $r-\phi$ view of the track segment in one chamber, where there are 8 measured points in most of cases, but cannot be used for the $r-z$ superlayers where only a maximum of 4 points are available. The drift velocity is estimated from the mean value of the track-by-track drift velocity distribution.

Fig.3.9 shows the mean values of the drift velocity for the MB2 chambers of wheel +2 using both methods. Mean timer values have large unexpected fluctuations from one sector to the other, related to the errors on the t_{trig} determination. These fluctuations are canceled when the fit method is used.

The average drift velocity values using the fit method for all the chambers are plotted in Fig.3.10, for runs without magnetic field and with the nominal magnetic field of 3.8 T.

The data at B = 0 T show an average value of 54.5 μm/ns for the drift velocity and the sigma indicates that differences between chambers are in the order of 0.2%. For B = 3.8 T a second peak is observed at values of 53.6 μm/ns, this peak corresponds to the MB1 chambers of the external

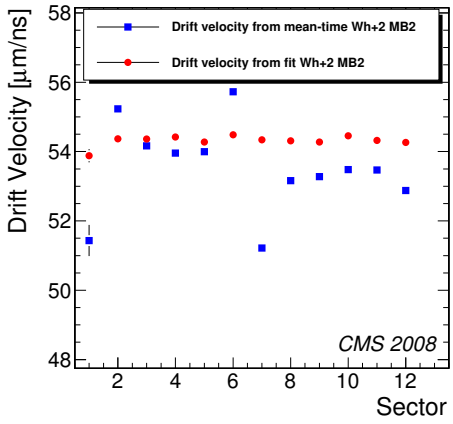


Figure 3.9: *Mean values of the drift velocity for the MB2 chambers of wheel+2 using the mean timer technique (blue square) and fit (red circles) methods.*

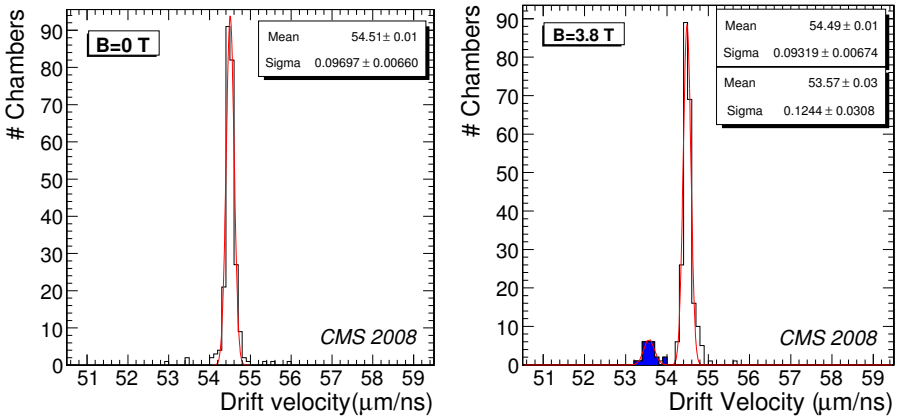


Figure 3.10: *Drift velocities values for $B = 0$ T (left) and $B = 3.8$ T (right). The small red peak corresponds to the MB1 chambers of wheels +2,-2 and shows the influence of a higher magnetic field in these regions.*

wheels and it is due to the presence of a higher radial magnetic field.

The effect of variation of the radial magnetic field along the z coordinate on the drift velocity (as calculated with the fit method) is shown in Fig.3.11. Positive wheels +1,+2 are not in the figure but they show the same behavior as wheels -1,-2, respectively. The presence of the magnetic field affects, as expected, only the MB1 chambers, primarily in the external wheels but some effects are also observed in wheels +1,-1.

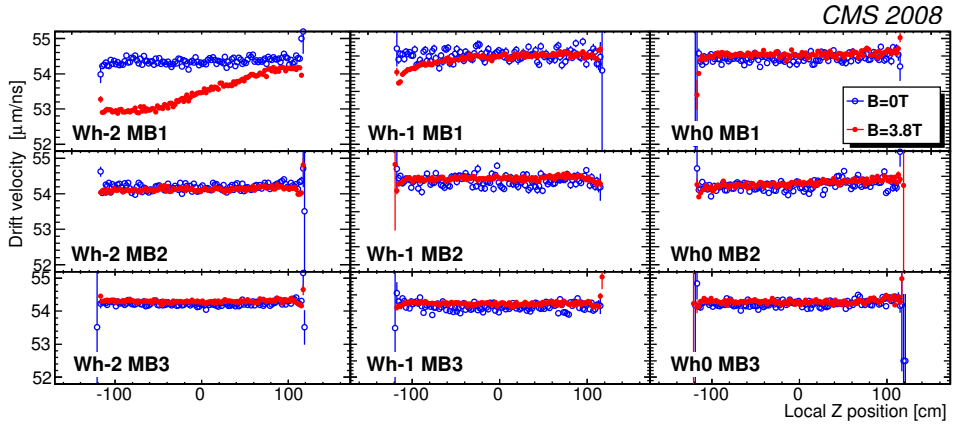


Figure 3.11: *Drift velocities as a function of the z position (measured by the $r-z$ superlayers), for $B = 0$ T and $B = 3.8$ T.*

A consequence of the more precise drift velocity calculation for cosmic data is the achievement of a better chamber resolution. The resolution obtained from the fast calibration procedure, described in this chapter, is of the order of 400/600 μm, as it is shown in Fig.3.5. This value is also confirmed taking into account the uncertainty of the slope of the rising edge of the time box (Fig.3.3) multiplied by the constant drift velocity of 54.3 μm/ns. A deatiled analysis on the chamber resolution obtained with the track fit method, described in [33], gives chamber resolutions close to the nominal value of 200 μm.

3.3 DT noise studies

Every hit registered by the DT front-end electronics with signal higher than the discriminating threshold of 30 mV and not associated to the passage of a particle, is considered as a noise hit. The DT random electronic noise is described by a flat distribution positioned at the bottom of the time box (Fig.3.12) For the DT calibration is crucial to take under control the noise behaviour because it can affect the fit to the rising edge of the time box and therefore the time pedestal computation. My contribution to the calibration studies focused mainly on the definition of a "noisy cell", the implementation of an automatic method to identify the total number of noisy cells and their position and on a detailed investigation of the noise geometrical distribution in the full DT system.

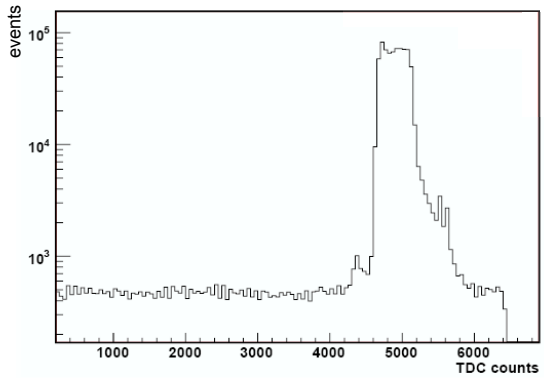


Figure 3.12: *TDC time distribution for a DT cell. The logarithmic scale on the y axis spotlight the presence of the flat distribution of noise hits in the lower part of the plot.*

During the first period of commissioning of the DT detectors, dedicated random trigger runs were taken for noise studies. Systematic analysis lead to define a cell as "noisy" if its rate is higher than 500 Hz. Random trigger runs permitted to study not only the average noise rate and its geometrical

distribution but also the time dependence.

The method used to investigate the noise of a DT cell as a function of time is based on the computation of the distribution of the time interval between two consecutive noise events. This distribution is fitted with an exponential parameterisation to extract the time constant. Fig.3.13 show the relation between this time constant and the average noise of the related cell for all the DT channels of wheel -2, sector 6. Cells with few noise hits have time constants concentrated around zero while for noisy cells there is a clear linear dependence between the time constant and their average noise. Entries in the lower part of the plot would be associated to cells which have become noisy after a certain time during the data taking, while entries in the higher part belong to cells which have rare noise events but with a very high number of hits per event. Cells with a discontinuous and high noise rate were often present during the first period of commissioning since the DT installation was on going. This phenomena gradually disappeared as the data taking conditions became more stable and the discriminating threshold for signal rised from 20 mV to 30 mV.

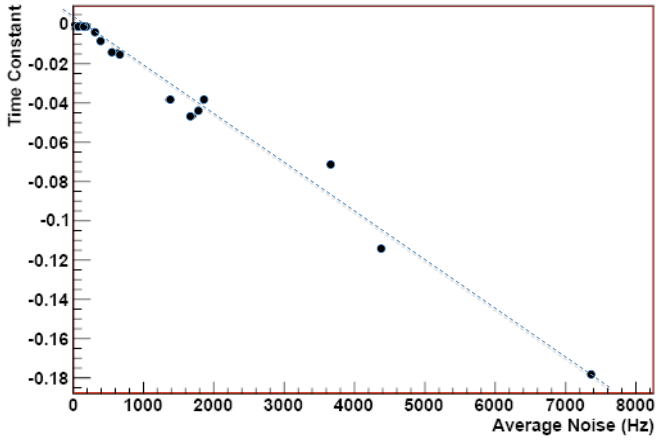


Figure 3.13: *Time dependence of the noise events in DT cells as a function of the related average noise rate.*

During the second period of DT chambers commissioning including the CRAFT data taking period, the noise analysis code had to be adapted for running on cosmic ray runs. Noise studies with cosmics runs have the advantage, with respect to the random trigger runs, of monitoring the detector in more realistic conditions. The idea was to perform the noise analysis selecting events in a time window of 350 ns before the Time Pedestal in order to avoid hits related to the passage of a particle through the detector. The DTs noise status found with this new method matches exactly with the results obtained using random trigger runs.

During CRAFT the first aim of the noise calibration was to check the stability of the number of noisy cells and the average noise rate in different conditions of the CMS detector: data samples with and without magnetic field and with different subdetectors in the data acquisition have been considered. For all the representative runs analyzed the number of noisy cells is around 0.01% of all the DT channels, as shown in Table 3.1. Moreover 50% of all the noisy cells keep their noisy status for all the runs. This means that the electronic noise involves few ten of cells in the full DT system and it does not depend on magnetic field or the type of subdetectors present in the data acquisition.

Run	Data	B (T)	Participating Detectors	# of noisy cells
57771	19/08/08	0	all, but CSC+PIX	18
66286	15/10/08	0	all, but CSC+PIX+RPC	24
67838	28/10/08	3.8	all, but CSC	30
68958	02/11/08	3.8	all	19
70195	09/11/08	3.8	all, but 14CSC+HO+DiskRPC	20

Table 3.1: *Summary of representative runs analyzed, the associated CMS data acquisition status and the number of noisy cells.*

It is also important to check the stability of the average noise rate of the drift cells, which is the mean value of the noise rate distribution (see Fig.3.14). For all the representative runs analysed it resulted $\sim$ 4 Hz.

This supports the previous considerations confirming the DT noise low and stable.

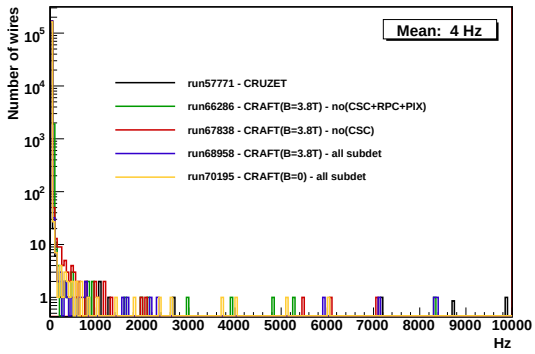


Figure 3.14: *Noise rate distribution measured for different conditions of data taking.*

Further studies concerning the noise geometrical distribution have been performed on the same representative runs. Fig.3.15 shows the distribution of noisy cells as a function of the DT station. The bulk of noisy channels is clearly located in the innermost chambers (MB1) where the internal cabling is more complex due to the reduced space.

In Fig.3.16 the noisy cells distribution as a function of the wire number is shown. Here, noisy cells appear concentrated in three main regions, due to the fact that different DT stations have a different number of channels. The first region corresponds to wires located into the left extremity of DT chambers while the second and the third ones are associated to the right extremities of stations (MB1: 47 wires, MB2: 59 wires, MB3: 73 wires, MB4: wires in the range 30-102). This mean that, for each DT station, noisy cells are concentrated at the layer extremities, where wires are close to the HV connectors.

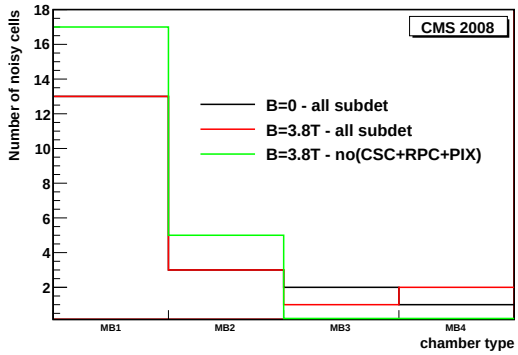


Figure 3.15: *Distribution of noisy cells as a function of the DT station.*

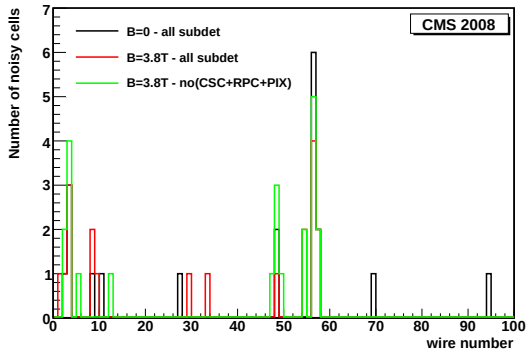


Figure 3.16: *Distribution of noisy cells as a function of the wire number.*

3.4 The DT calibration workflow

A fast calibration of the DT system is crucial to feed prompt data reconstruction with fresh calibration constants. During collision and cosmic ray data taking periods, the calibration parameters had to be produced, validated and made available to be used in reconstruction within one day from

the data taking.

The number of calibration regions has to be a compromise between the necessity of keeping things simple, not requiring too high statistics, and the need of reducing systematic errors by separately calibrating regions where parameters may take very different values. As mentioned in previous sections, the superlayer granularity has been identified to be the most suitable calibration unit. In order to reach the precision obtainable with the fast calibration, about 10^4 tracks crossing each superlayer are required.

The DT calibration processes are part of the central CMS calibration and alignment work-flow, which is described in details here [34] and illustrated in Fig.3.17.

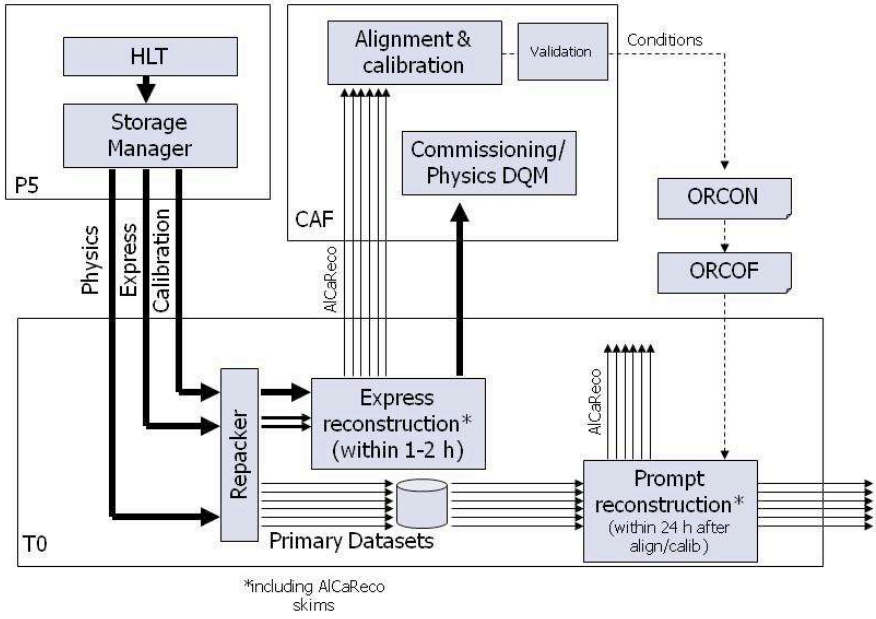


Figure 3.17: *Prompt Calibration work-flow for the CMS Calibration and Alignment processes.*

The main steps of the calibration data work-flow may be summarized

as follow:

- the online processing (High Level Trigger process) at the CMS detector site, the so-called *Point Five* (P5);
- the offline processing at the central CMS computing centre (T0) including special alignement and calibration dataset, called *AlcaReco* samples;
- the offline processing at the CERN Analysis Facility (CAF) where the Condition Data Bases are produced and validated.

The *AlCaReco* sample contains a reduced number of events and data content just to fulfill the requirements of the calibration tasks. The sample is saved to the CAF and it is taken as input to the calibration process. The calibration algorithm runs on the CAF and produces a set of constants, usually in the form of a temporary database. The constants are then submitted to a validation procedure before being copied to the central CMS ConditionsDatabase (ORCON database). Finally, they are automatically streamed to the ORCOF database and are available to the CMS prompt reconstruction.

The quality and stability of the calibration constants is a crucial part of the procedure and must be continuously monitored. I've devoted part of my work implementing these validation procedures within the central CMS Data Quality Monitoring (DQM) framework [35], described in Section 4.3.1. Data quality assessment for the DT calibration constants consists mainly in:

- the definition of the acceptance criteria used to validate the constants. To satisfy this requirement, dedicated quality tests are applied to the residual distributions calculated at the different steps of the calibration work-flow and periodically checked by shifters.
- the monitoring of time stability of the calibration constants and the search for continuous trends or sudden changes in operating conditions. For this task a special tool have been implemented to compare

the currently produced database with a reference database and to give a report on the stability of the calibration constant under study. For the noise database, which is supposed to change more than the other constants, a dedicated analysis is provided. First a general comparison of the total number of noisy cells between the reference and the new database is performed. Then a geometrical study of the noisy cells found in the new databases is done in order to make easier the search for new sources of noise inside the DT system.

All the calibration constants described in the present chapter have their validation and monitoring processes, and for each of them detailed and summary DQM plots are provided. The DT condition constants have been monitored through the entire CRAFT data taking period and have shown generally a good stability in time. An example screenshot of the DQM web interface is shown in Fig.3.18. The plots show the quality test results on the mean of residuals. The green boxes are associated to good results, the yellow and the red ones correspond to warning and alarm messages, respectively.

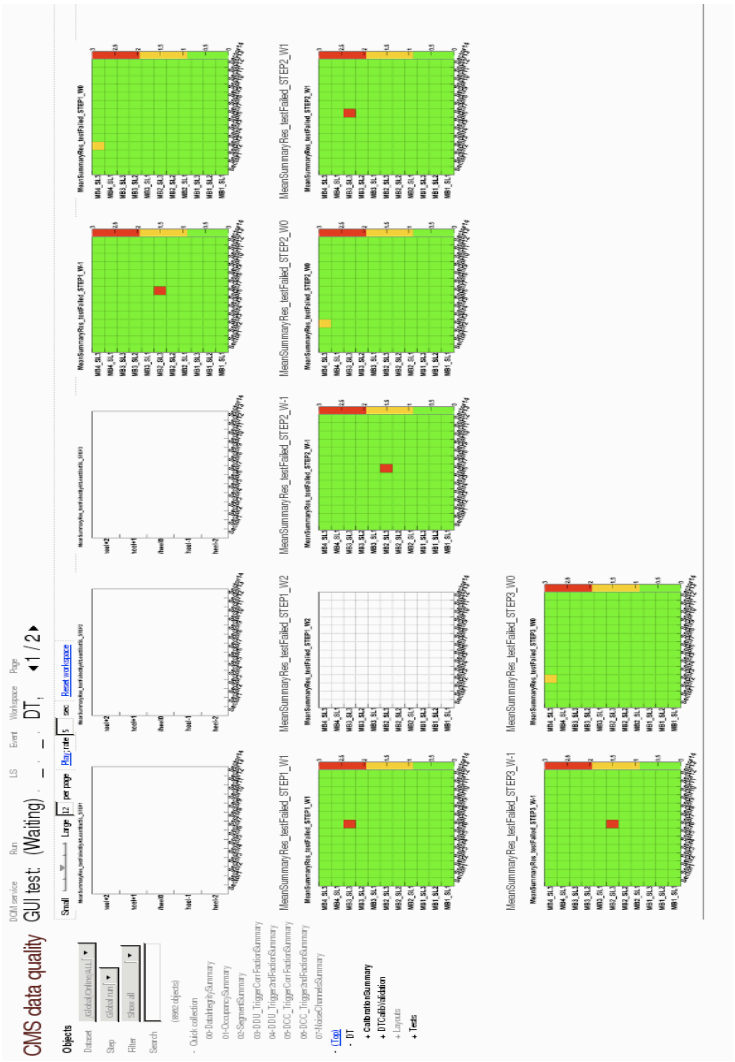


Figure 3.18: *Example of a web page used by the offline DT shifters to monitor the calibration parameters.*

Chapter 4

Local Reconstruction in the Drift Tubes

The DT local reconstruction is the fundamental link between the output of the Data Acquisition System and the input parameters required to the DT chambers for the fit of the muon track inside the full CMS spectrometer. The procedure, the performance and the online monitoring of the local reconstruction during CRAFT are described in this chapter.

Local reconstruction begins with the computation of mono-dimensional hits in the single drift cells. The only information contained in these hits is their distance from the anode, with an intrinsic left/right ambiguity and without any information about their position along the wire. Section 4.1 describes in details this first reconstruction step.

In Section 4.2 the segments building, which starts from the hits recorded in each single cell to reconstruct first the $r-\phi$ and $r-z$ segment projections separately, is presented. These two-dimensional segments still do not provide any information about the coordinate along the direction of the wire, but they measure the track angle in the measurement plane (orthogonal to the wires). Combining the two projections it is possible to reconstruct the direction and position of the muon crossing the chamber. The resulting three-dimensional segments are the input to the muon track fit in the CMS

spectrometer.

The online monitoring of the DT local reconstruction together with the main aims of the online Data Quality Monitoring of CMS are presented in Section 4.3.

4.1 Reconstruction of hits position within cells

The primary objects that result from the DT local reconstruction are points in the cell volume, called "RecHits". These objects are built computing the drift distance corresponding to the measured drift-time. This requires the knowledge of the average drift velocity in the cell, considering also its dependence on the residual magnetic field and on the track angle. As already discussed, two reconstruction algorithms have been developed on the CMS software: one uses a constant drift velocity over the entire cell, the other is based on a time-to-distance parameterization obtained using a GARFIELD [29] simulation of the cell behaviour. This latter method takes into account the dependence of the drift velocity and of the cell non-linearities on the track angle and on the residual magnetic field, thus achieving better resolution. The parameterized drift velocity algorithm has been improved on $p-p$ simulated data and could not be used with cosmic data because the muons, which have a big impact angle with respect to the normal of the chamber, generate an ionization shower with different properties. For this reason only the method based on the constant drift velocity is described in Section 4.1.1.

As discussed in Section 3.2.1, corrections to the synchronization constants due to the time-of-flight and the signal propagation along the wire must be applied in order to achieve the best possible resolution. These corrections, as well as the parameters used as input for the cell parameterization (the track incidence angle and the magnetic field) cannot be computed for an individual TDC measurement, which only contains information about the hit distance from the wire. For this reason an iterative reconstruction procedure is adopted:

- *first step*: reconstruction at the cell level: the left-right ambiguity is

unsolved and the position of the hit along the wire is not yet determined;

- *second step*: the hits are used to build a segment within a superlayer. The left-right ambiguity is solved. Using the cell parameterization and the impact angle, which is then known, the hit position can be recomputed;

- *third step*: the hit is used to fit a three-dimentional segment with both $r-\phi$ and $r-z$ projections. At this point, also the position along the wire is determined and can be used for a more precise estimation of the residual magnetic field which may change along the tube. Also the synchronization offsets due to the time-of-flight and the signal propagation along the wire can be refined, as discussed in Section 3.2.1.2, so that hit position can be further improved.

To summarize, the reconstruction at the cell level gives the input hits used to build a segment. When a first segment is reconstructed, the positions of its hits are further refined and the hits are then refitted. In order to discard noise and pile-up signals, the reconstruction is performed only for drift times falling in a user-defined time window, while the remaining TDC measurements are discarded.

4.1.1 The constant drift velocity method

The good uniformity of the electric field within the cell allows the reconstruction of the distance from the wire assuming a constant drift velocity over the entire cell:

$$x = v_{drift} \cdot t_{drift} \tag{4.1}$$

The constant drift velocity algorithm does not account for local variations of the magnetic field in the tube nor for different impact angles of the tracks. This method relies on the calibration procedure described in Sec.3.2.2.2, which can compute the average drift velocity which best fits the working

conditions of a group of cells. Currently, the calibration is performed separately for each superlayer, so that the resulting drift velocity corresponds to the average conditions in each superlayer.

Such simple reconstruction algorithm allows satisfactory results especially for tracks with small impact angles and in regions with low residual magnetic field, as in these cases the cell non-linearities are less important.

For this algorithm, the three-step reconstruction procedure is only used for the refinement of the synchronization constants.

4.1.2 Results with cosmic data

As already discussed in Section 3.2.3, the distribution of the residuals of reconstructed hits with respect to the position predicted from the extrapolation from track segments is largely dominated by the uncertainty on the muon arrival time inside the 25 ns time window associated to the first level trigger. Therefore, after the local pattern recognition has been performed, the arrival time of the muon can be treated as a free parameter in the refitting of the segment in order to determine the final segment position and direction. The results of an accurate study [33] which makes use of this method with on CRAFT data are shown in the following.

RecHit resolution The distribution of the reconstructed hits residuals after the refitting procedure is shown in Fig.4.1 for sector 4 of the external wheel -2. Here, the residual magnetic field in the DTs volume has the largest variation along the chambers length, reaching the highest values (up to 0.6 T for the radial component in the MB1 stations). The scatter plots of the residuals as a function of the local position of the hit in the DT cells are also shown in Fig.4.1, displaying the good linearity of the cell behaviour in the whole drift volume.

The distribution of the hit resolution, obtained by the fit to the residuals distribution, in all the DT chambers is shown in Fig.4.2. The blue entries in the histogram are associated to vertical chambers. For most of the stations, the resolution is ~ 300 μm.

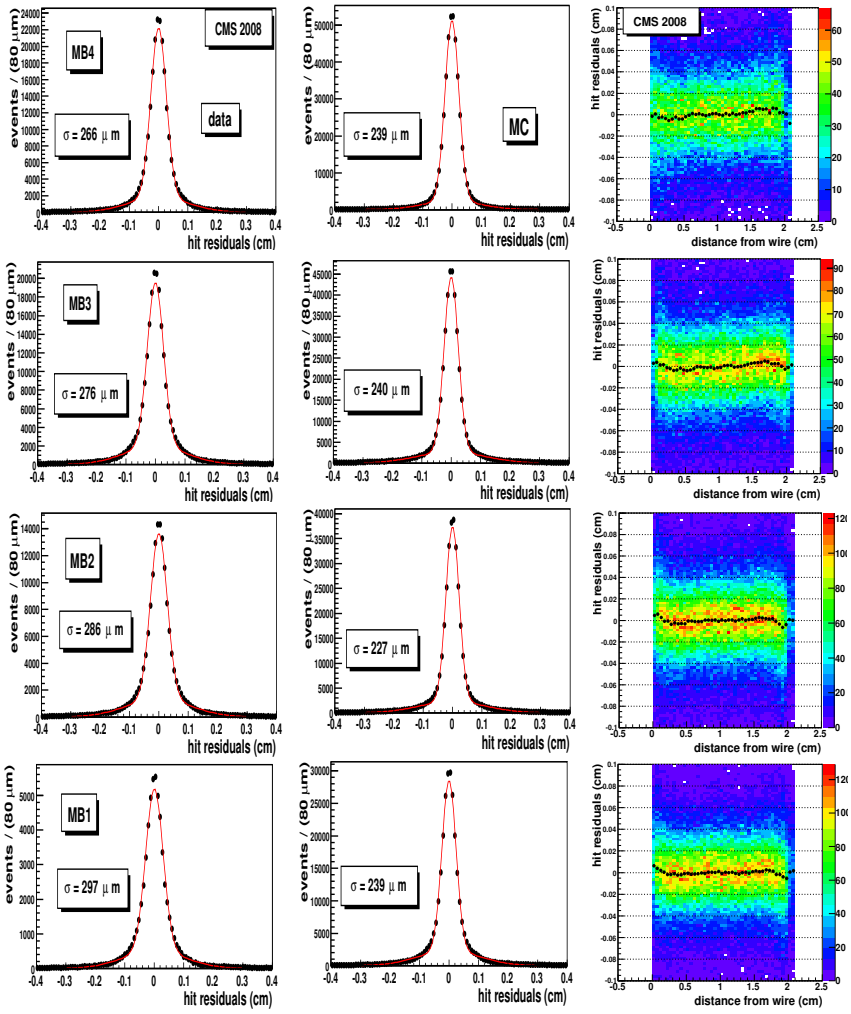


Figure 4.1: *Residuals of reconstructed hits in four DT chambers, atfer the fit to the muon arrival time. Left column plots: data; Middle column: MonteCarlo simulation; Right column: scatter plot of residuals as respect to the hit position in a DT cell.*

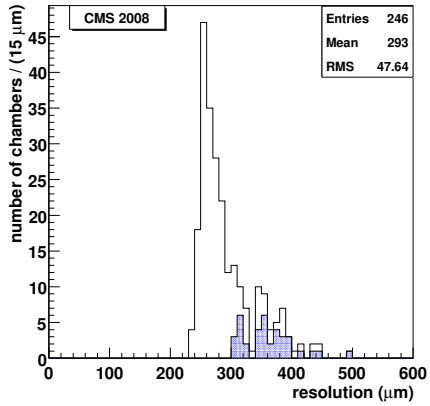


Figure 4.2: *Distribution of the RMS values of the fit to the reconstructed hit residuals in all DT chambers, after the segment refit. The blue entries are from stations in the vertical sectors.*

RecHit efficiency. The hit reconstruction efficiency is computed in three consecutive steps:

- reconstruction of a three-dimensional DT track segment excluding the hits in the relevant layer;
- extrapolation of the the track segment to the relevant layer;
- search for the presence of a reconstructed hit in the extrapolated cell.

Fig.4.3 shows the hit reconstruction efficiency as a function of the predicted hit position in the cell for the MB1 stations. The efficiency drop near the anode wire position (x=0 in the plots) is due to the pedestal subtraction procedure described in the previous chapter, and is well reproduced in the Monte Carlo simulation. No significant differences between the data at B=0 and B=3.8 T is observed.

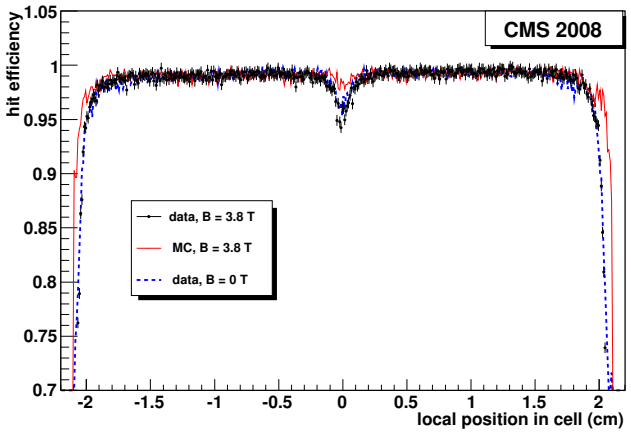


Figure 4.3: *Efficiency to have reconstructed a hit in a cell crossed by a cosmic muon, as a function of the predicted muon position in the cell, for the MB1 stations. The $x = 0$ position corresponds to the location of the anode wire in the cell.*

4.2 Segment Reconstruction

One of the main aims of the segment building procedure is to solve the left-right ambiguity of the mono-dimensional hits and to provide track stubs, thus helping the pattern recognition during the muon track reconstruction (Chapter 5). Moreover the segment reconstruction provides a precise determination of the track impact angle and of the hit position along the wire, which can be used to refine the computation of the drift distance (cf. Section 3.2.2.2).

The segment reconstruction acts in each chamber on the $r-\phi$ and $r-z$ projections independently. At the end of the procedure the two projections are combined and three-dimensional segments are built. The algorithm performs in each projection the pattern recognition and a linear fit of the hits. A linear fitting model, given the hit resolution and the small height of a chamber, is a correct approximation.

The reconstruction is performed in three steps:

- segment candidates are built from sets of aligned hits;
- the best segments among those sharing hits are selected, solving conflicts and suppressing ghosts;
- the position of the used hits is then updated using the information from the segment and the segments themselves are re-fitted.

4.2.1 Segment building: the procedure

Search for segment candidates. The reconstruction begins identifying pairs of hits in different layers, starting from the most separated layers. A pair of hits is kept if the angle of the proto-segment is compatible with a track pointing to the nominal interaction point, within a configurable tolerance. The tolerance is usually set to 0.1 rad in the $r - z$ projection and 1.0 rad in $r - \phi$ plane (as this is the projection interested by the track bending). These constraints can be switched off in the reconstruction of cosmic muons. As each hit has a left-right ambiguity, both hypotheses are considered if they fulfil the above condition.

For each pair, additional compatible hits are searched for in all layers. Then the building algorithm checks if the left or the right hypotheses have a distance from the intercept smaller than 10 times the error on their position. It is possible that both the left and right hypotheses are compatible with the segment. In this case, both candidates are retained and the ambiguity is solved later. Also, a muon can cross the I-beam separating two cells generating a signal in both. In this case, both hits are considered for that layer, in order to avoid any bias.

Once the pattern recognition is completed, each collection of hits is fitted using the positions and the errors assigned to the hits. For each pair of hits only the segment candidate with the highest number of hits and the smallest χ^2 is retained; all the others are rejected. Finally, a loose quality criterion is applied, requiring the number of hits ≥ 3 and the $\tilde{\chi^2} < 20$ for all the candidates.

This algorithm is applied directly to each $r-z$ superlayer. It can also be applied to the individual $r-\phi$ superlayers to make performance studies, however for track-fitting purposes it is more convenient to build the $r-\phi$ projection using the hits from both the $r-\phi$ superlayers in the chamber.

It should be noted that the procedure described above is not attempted if the number of hits in a given projection is larger than a programmable number (default is 50). This is required in order to avoid particle showers which can generate a very high occupancy in the chamber. A so great amount of hits increases the combinatorics deteriorating the performance of the pattern recognition. Therefore a complementary algorithm to treat such cases is under study.

Segment selection. Once the pattern recognition is done, a consistency check is performed in order to test whether two candidates use the same hits but with different left-right hypotheses. In that case, the conflicting hit is removed from the worst of the two segments, where the quality is defined by the number of hits and the χ^2 of the segment.

Moreover two segment candidates can share all their hits but with different left-right hypotheses for each hits. The two candidates have in this case the same quality but different incidence angles. In $p-p$ collisions, the choice is implicit in the request of compatibility with the nominal interaction point described previously. When this requirement is removed, as for the reconstruction of cosmic muons, two options are available: either retaining both candidates, thus leaving the muon track fit algorithm to select the best segments, or choosing the candidate with the smallest angle with respect to the vertical to the chamber among the two ones.

As the incidence angle of the segment is now known, the position of the hits in the remaining segment candidates are updated. This is the so-called "second step", mentioned in Section 4.1. The segment linear fit is recomputed using the updated hits.

Matching of the two projections. Finally, the $r-\phi$ and $r-z$ projections are handled independently to reconstruct a three-dimensional segment. As the two projections are orthogonal, a segment in one projec-

tion cannot be used to validate or invalidate a segment in the other: all combinations of segments from the two projections are kept.

The additional knowledge of the position along the wire is than used to update the hit position in the cells (the "third step" described in Section 4.1) before performing the final fit of the segment. The result is a segment inside a chamber suitable for use in the track reconstruction (Chapter 5).

4.2.2 Results with cosmic data

Hits multiplicity. The first quantity directly related to the segment quality is the multiplicity of the associated hits. The distribution of the hit multiplicity is shown in Fig.4.4, for muon tracks passing at a distance larger than 1 mm from the I-beams, thus excluding the small regions where the agreement between data and simulation on the hit reconstruction efficiency is poor. The distributions are peaked, as expected, at the total number of layers in the chamber (8 in MB4 and 12 in the other stations), although the Monte Carlo simulation shows a slightly bigger average value.

The distribution of the segment incident angle with respect to the vertical axis in the bending plane is shown in the right column plots of Fig.4.4 The cosmic data distribution is well reproduced by the simulation. The observed increase of the spread around the normal direction when passing from MB4 to MB1 is due to the bending effect of the magnetic field in the iron yokes.

δ-ray studies. The discrepancy between data and simulation in the associated hit multiplicity distributions is also related to the efficiency of reconstructed hit association, affected by the occurrence of δ-ray electrons originating in the gas volume and in the mechanical structure of the chamber. If these electrons pass closer than the original muon to the anode wire of the cell, they mask the following muon signal if it arrives within the electronics dead time of 150 ns. Fig.4.5 shows the distribution of the difference between the distance from the cell anode wire of the first hit recorded (independently from its association to the muon track segment) and the distance of the position of the track extrapolation in the cell. The

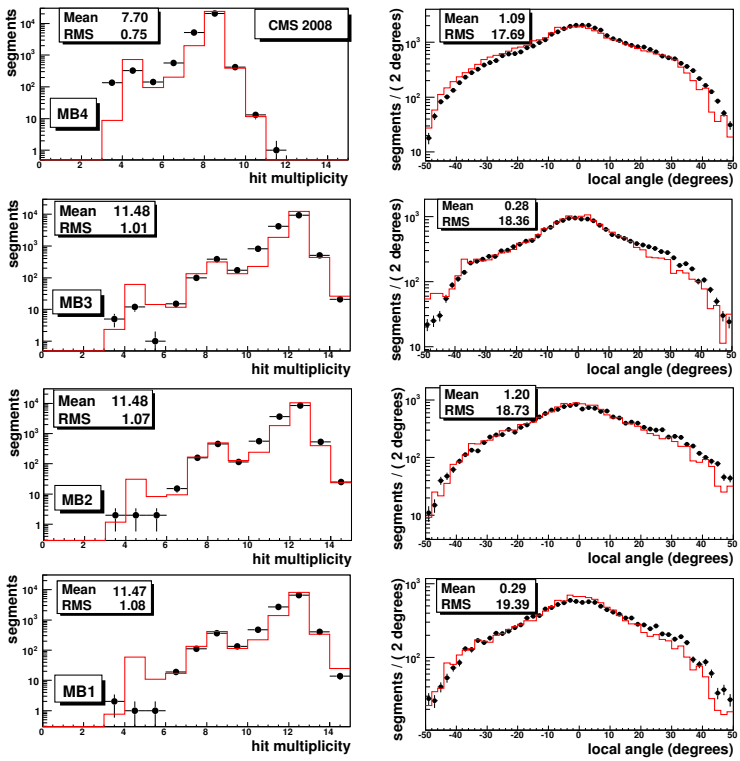


Figure 4.4: *Multiplicity of associated hits in reconstructed segments in sector 4 of wheel -1 (L). Segment direction as respect to the vertical axis (R). Real data (points) and simulated data (solid line histogram) are shown.*

population at large values of the distance difference is due to the δ-ray hits that are not associated to the track segment. The tail at positive values of the difference is due to events with a δ-ray where the muon hit goes undetected. Moreover this tail has been extended to values bigger than the half-cell dimension to show the population from neighbouring cell in the same layer. The data and simulation distributions show a reasonably good agreement, both in the absolute yield of δ-rays and in the asymmetry of

the distribution, with however a slight underestimation of the effect in the simulated data.

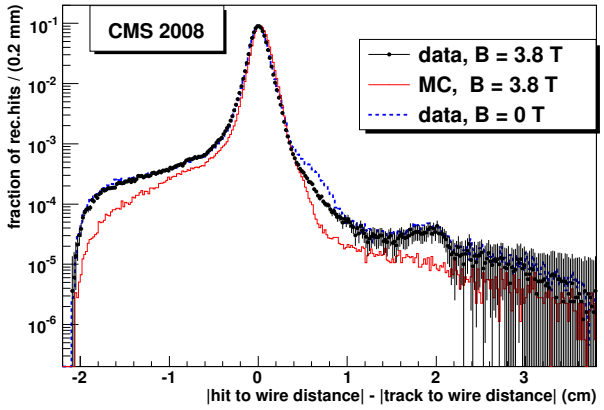


Figure 4.5: *Distribution of the difference between the distance to the cell anode wire of the first hit recorded in a cell and the distance of the extrapolated track position.*

Segment reconstruction efficiency. The muon tracks independently reconstructed in the silicon Tracker can be extrapolated to the DT chambers in order to evaluate the segment reconstruction efficiency. Only tracks with $p_T > 30$ GeV/c with a space window of 20 cm around the predicted position (which takes into account the observed effect of the multiple scattering) have been considered for the efficiency computation. Moreover to ensure a reliable extrapolation from the Tracker tracks, when computing the efficiency for a given chamber MBn, the extrapolation of the track to station MBn+1 was required to be confirmed by a good DT segment also in this chamber, within the same acceptance window defined above.

The method can be safely applied to all chambers of the three uppermost and lowermost sectors of the wheels YB-1, YB0 and YB+1, where there are enough Tracker tracks with good quality and reliable extrapolations. The DT chamber efficiency can be also evaluated making use exclusively of

the information coming from the muon spectrometer, thus extending the efficiency measurement to the chambers of all the barrel wheels. Muon tracks are reconstructed with the information provided by neighbouring chambers and extrapolated to the mid-height of the chamber under test. In the central wheels, the results are found fully compatible with those obtained using the extrapolation from the inner Tracker.

All DT sectors except the vertical ones (S1 and S7) of the five wheels have been studied. Figure 4.6 shows the chamber efficiency in the $r-\phi$ plane. Every plot contains the efficiencies for a given station (MBn) for each sector and wheel analyzed, marked on the horizonal axis. Figure 4.7 shows the corresponding efficiencies in the $r-z$ plane. Results on efficiencies are fully compatible among sectors; the drop of efficiency observed in some of them corresponds to sectors where the muon incident angle is largest and the segment reconstruction is worse.

Bending power measurements. Data with a magnetic field of 3.8 T in the central solenoid were considered to study the bending power of the muon spectrometer. The distributions of the difference of the track angle measurements between consecutive stations were studied for different values of the transverse momentum of the associated track as independently measured by the Tracker detector.

These distributions are shown in Fig. 4.8 for MB2-MB3 pairs of stations. As seen from the figure, the bending power for a $p_T = 30$ GeV/c muon is about 6.6 mrad. Similar distributions are observed for MB1-MB2 and MB3-MB4 pairs of stations, with bending power equal to 4.0 mrad and 6.0 mrad, respectively. It is worth to remind that the width of the magnetized steel between the chambers is about 30 cm between MB1 and MB2 and 62 cm between MB2-MB3 and MB3-MB4. The magnetic flux density in the steel yokes decreases slightly with the radial position.

Figure 4.9 shows the distributions of the angle difference between MB1 and MB4 stations, displaying the bending power of the full lever arm in a barrel sector. For muons selected in the p_T range [150, 250] GeV/c, the average deflection by the magnetic field in the steel return yokes of the magnet is about 3.4 mrad.

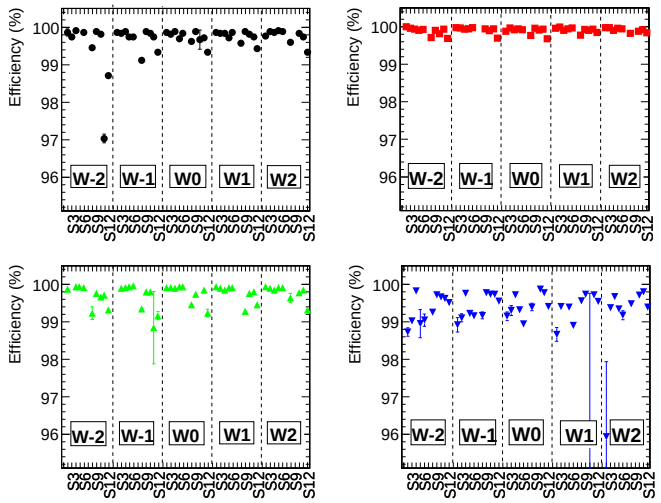


Figure 4.6: *Segment reconstruction efficiency in $r-\phi$ plane in the Muon Barrel chambers: Top plots: MB1 (left) and MB2 (right) chambers; bottom plots: MB3 (left) and MB4 (right) chambers.*

4.3 The DT online Data Quality Monitoring

In this section, a brief overview of the main aims of the DQM (Sec.4.3.1) of CMS and in particular of the online system (Sec.4.3.2) is done. Then what is actually monitored for the DT local reconstruction is presented in order to check the quality of the three-dimensional segments during the data taking (Sec.4.3.3).

4.3.1 The CMS Data Quality Monitoring

Data quality monitoring is critically important for the detector and operation efficiency, and for the reliable certification of the recorded data for physics analyses. The CMS experiment has standardised it on a single end-to-end DQM chain (Fig.4.10). The system comprises:

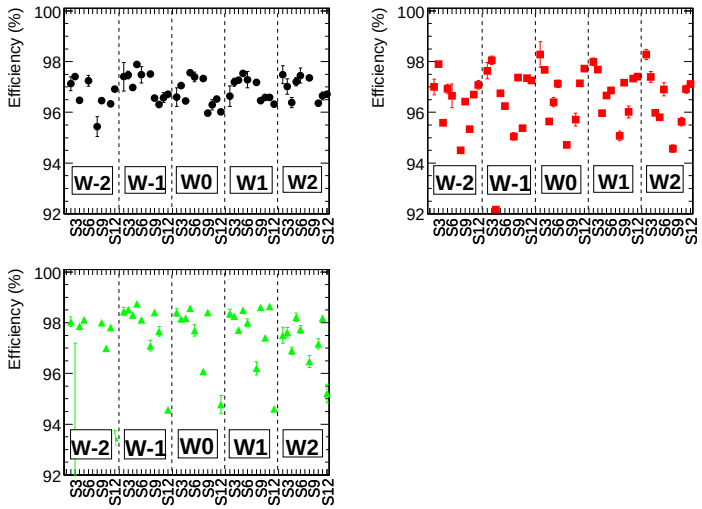


Figure 4.7: *Segment reconstruction efficiency in $R-z$ plane in the Muon Barrel chambers: Top plots: MB1 (left) and MB2 (right) chambers; bottom plot: MB3 chambers.*

- tools for the creation, filling, transport and archival of histogram, with dedicated algorithms for performing automated quality and validity tests on value distributions;
- online systems for the monitoring of the detector, the trigger, the DAQ hardware status and data throughput and also for the online reconstruction and the validation of the calibration results, software releases and simulated data;
- visualisation of the monitoring results;
- certification of datasets and subsets for physics analyses;
- retrieval of DQM quantities from the conditions database;

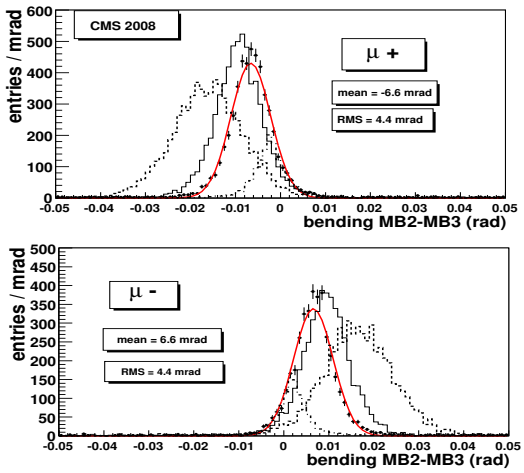


Figure 4.8: *Bending angle differences between MB2 and MB3 stations. Top: μ^+; bottom: μ^-. Distributions for different p_T intervals are shown: [8-12] (dashed line), [18-22] (full line), [27- 33] (points) and [90-110] GeV/c (dashed-dotted line). The curves show the result of a Gaussian fit to data distribution for the $150 < p_T < 250$ GeV/c sample.*

- standardisation and integration of the DQM components in CMS software releases;
- organisation the activities, including shifts, tutorials and training operations.

The high-level goal of the system is to discover and exploit errors and general problems occurring in detector hardware or reconstruction software, with sufficient accuracy and clarity to reach good detector and operation efficiency. Toward this end, standardised high-level tools distil the body of quality information into summaries with significant explaining power. The DQM activities are divided in online and offline operations for data processing, visualisation, certification and sign-off, as illustrated in Fig.4.10. Moreover the DQM supports mostly automated processes, but the usage

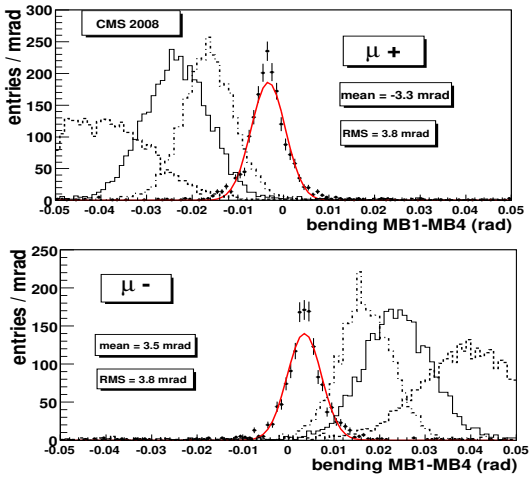


Figure 4.9: *Bending angle differences between MB1 and MB4 stations. Top: μ^+; bottom: μ^-. Distributions for different p_T intervals are shown: [8-12] (dashed line), [18-22] (full line), [27- 33] (dashed-dotted line) and [150-250] GeV/c (points). The curves show the result of a Gaussian fit to data distribution for the $150 < p_T < 250$ GeV/c sample.*

of the tools is also foreseen for the interactive and semi-automated data processing at the CAF analysis facility (see Section 3.4).

4.3.2 The online DQM system

4.3.2.1 Data processing

As illustrated in Fig.4.11, the online DQM applications are an integral part of the rest of the event data processing at the CMS detector site (Point-5). DQM distributions are created at two different levels: the high-level trigger filter units and the data quality monitoring applications.

The high-level trigger filter units process events at up to 100 kHz and produce a limited number of histograms. The histogram monitor elements are delivered from the filter units to the storage managers at the end of

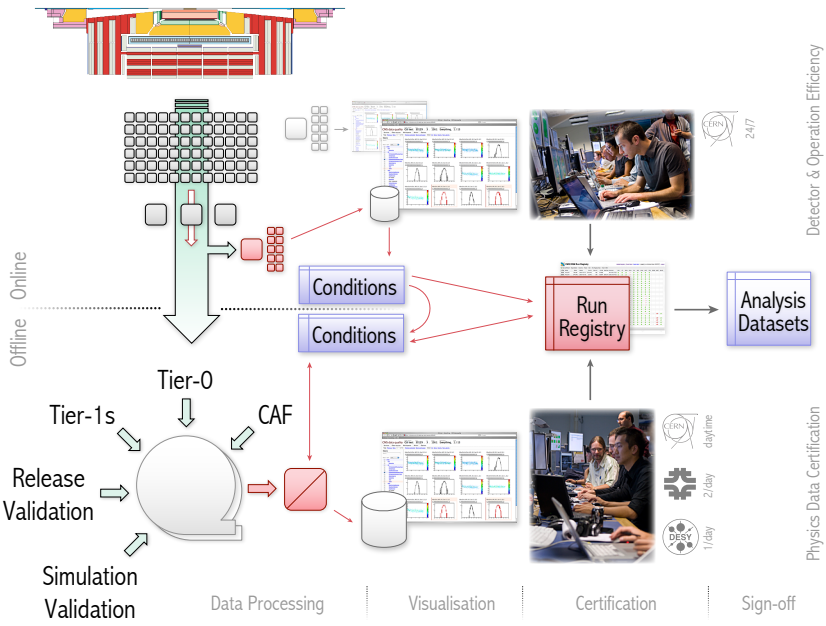


Figure 4.10: *DQM system overview.*

each luminosity section. Identical histograms across different filter units are summed together and sent to a storage manager server, which saves the histograms to files and serves them to DQM consumer applications together with the events data.

The data quality monitoring applications receive event data and trigger histograms from a DQM event stream from the storage manager at the rate of about 10-15 Hz. Usually one application per subsystem is present. Events are filtered for the stream by applying trigger path selections specified by the DQM group. Each DQM application requests data specifying a subset of those paths as a further filter. The DQM stream provides raw data products only, and on explicit request additional high level trigger information. Each application runs its choice of algorithms and analysis modules

and generates the results in the form of monitoring elements. The applications re-run reconstruction according to the monitoring needs. The monitor element output includes reference histograms and quality test results. The latter are defined using a generic standard quality testing module, set by configuration files.

4.3.2.2 Visualization

All the resulting monitor elements, including alarm states based on quality test results, are made available to a central DQM GUI for visualisation in real time, and stored to a ROOT file from time to time during the run. At the end of the run the final archived results are uploaded to a large disk pool on the central GUI. There the files are merged to larger size and backed up to tape. The automatic certification summary from the online DQM step is extracted and uploaded to a run registry.

4.3.2.3 Operation

Detector performance groups provide the configuration applications to execute, with the choice of conditions, reference histograms and the quality test parameters to use and any code updates required. Reviewed configurations are deployed into a central playback integration system, where they are first tested against recent data for about 24 hours. If no problems appear, the production configuration is upgraded. This practice allows the CMS detector to maintain high quality standard with reasonable response time.

4.3.3 The online monitoring of the DT local reconstruction

As explained above, to test the quality of the local reconstruction during the data taking, the DT DQM application need to re-run the segment building starting from the raw data which come from the DQM stream. My work focused on the implementation of a list of histograms to investigate the segment properties and check their distributions. The monitoring should be simple, synthetic and with an high efficiency in spotting problems. This because the main requirement for the DQM applications is to produce the

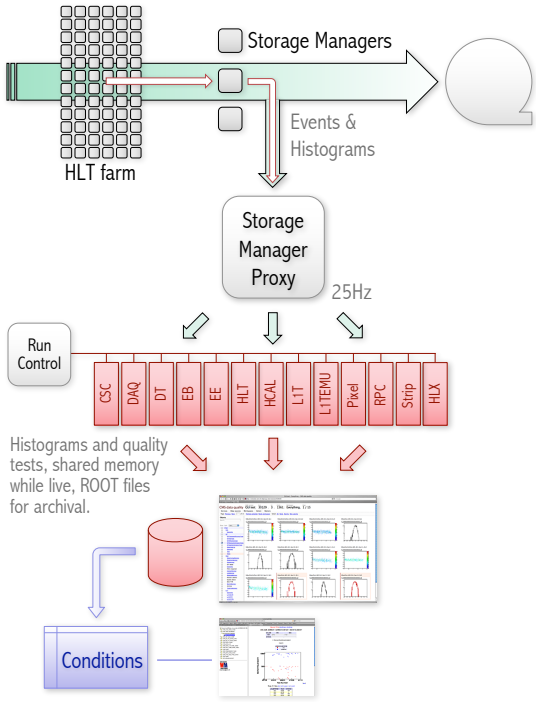


Figure 4.11: *Online DQM system.*

smallest number of histograms necessary to investigate the status of the parameters of interest and with the lowest operating costs in term of CPU memory and processing time.

The DT online Data Quality Monitoring at present make a control on different quantities as the data integrity, the hit time arrival distributions, the hit occupancy, the segment quality and finally the local cosmic and test-pulses trigger performances. In particular, to check the DT local reconstruction, the DT DQM produces and tests, for each station, the following histograms:

- the number of reconstructed segments per event;

- the number of hits per segment. Quality tests require this distribution to peak at 8 for the MB4 stations and at 12 for the other stations;
- the residuals between reconstructed hits and their associated segments. To test these distributions, the residuals are fitted with a gaussian function, whose mean is required to be compatible with 0 within a defined range (set by configuration file).

Results of all the tests described above are collected in summary histograms, shown in Fig.4.12. The information which these histograms contain is very simple: the bins are green if both the tests are good, while they are yellow if only the test on number of hits or on residuals have succeeded and red if both tests have failed. This summary is done with different level of granularity, for both stations and sectors. This except for the first global summary plot on the upper-left part of the figure which is green if at least 30% of the stations in each sector gives good test results. In this way central shifter can have a quick and clear overview of the DT segment status during the data taking with a refresh of the quality tests results at each luminosity section.

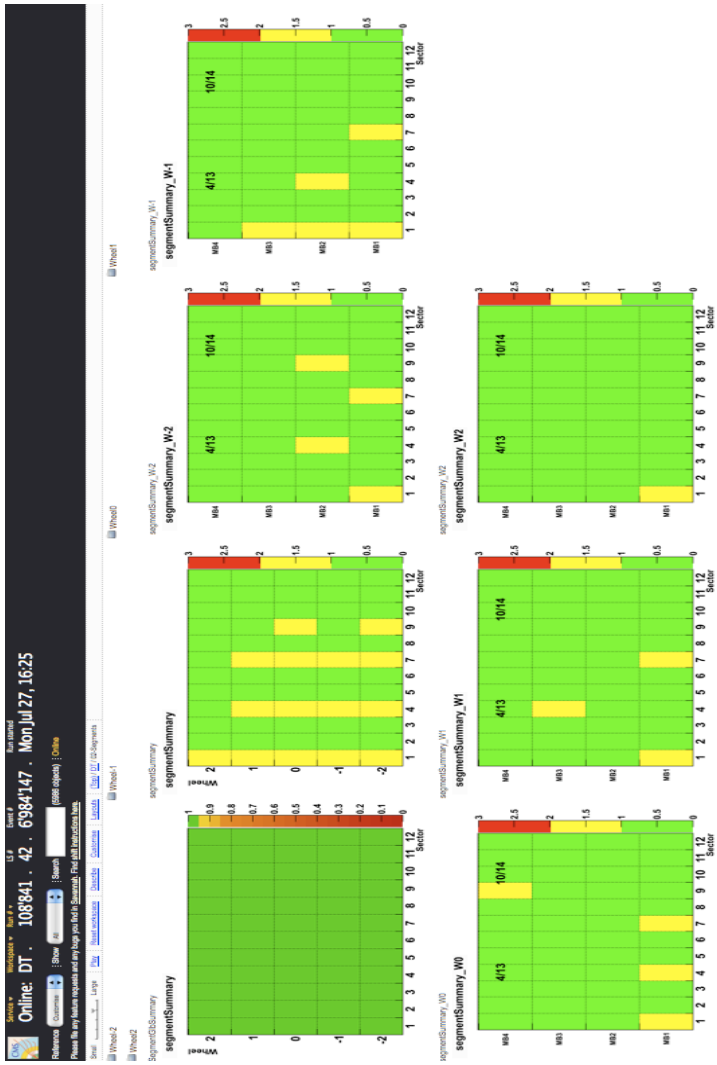


Figure 4.12: *Web interface example of the online DQM for the DT local reconstruction.*

Chapter 5

The CMS Muon reconstruction

The CMS muon reconstruction, which is a crucial point for the proper recongnition of various physics signatures, is divided in two main parts: a first fit is performed using hits coming only from the muon spectrometer, then muon tracks are propagated into the silicon tracker and the information from both systems, together with the muon energy deposits in the calorimeters, are combined. These steps, called *Stand-Alone Muon Reconstruction* and *Global Muon Reconstruction*, are both based on the Kalman Filter technique (Sec.5.1) and explained in Sec.5.2 and Sec.5.3, respectively. A brief description of the muon identification is in Sec.5.4, while the performance of the muon reconstruction with CRAFT data, its monitoring and validation are discussed in Sec.5.5 and Sec.5.6.

5.1 The Kalman Filter algorithm

The Kalman Filter technique [36], the basis of the CMS muon reconstruction, is a recursive method for the fit of a discrete set of data. The basic problem of this method consists in the estimation of a generic state vector

$\mathbf{x}$ given a set of measurements $\mathbf{m}_k$ that are assumed to have the form

$$m_k = H_k x_{k,true} + \epsilon_k \tag{5.1}$$

where H_k is the transform matrix from the state space to the measurement space and ϵ_k is the noise that affects the true state. In the case of the track reconstruction, the state vector is defined as the position and momentum relative to a given surface:

$$x = \begin{pmatrix} q/p \\ tan\phi \\ tan\theta \\ x \\ y \end{pmatrix} \tag{5.2}$$

where q is the charge, p is the momentum and ϕ, θ, x and y identify the track direction and position on the surface.
To complete the track description, a 5D curvilinear covariance matrix, the number of degrees of freedom and a summary of the information on the collected hits are stored in the corresponding data format.
The first step of the track reconstruction consists in the seed state estimation. For muon tracks the seed can be estimated starting from the measurements themselves or from external input. Each Kalman Filter step is then developed by two basic components:

1. *the propagator*: it extrapolates a state vector and its covariance matrix in a non-constant magnetic field, taking into account the effect of energy loss and multiple scattering in the material traversed by the track.

2. *the updator*: it includes the information from a measurement into the track.

The first result of a Kalman filter is a state on the surface of the last measurement, which includes all available information. However, the trajectory parameters calculated at other points of the trajectory do not include the

information from all measurements. A second iteration is used to update the parameters at every surface. In the Kalman filter terminology, this procedure is called *smoothing*. Once the hits are fitted and the fake trajectories removed, the remaining tracks are extrapolated to the point of closest approach to the beam line. In order to improve the p_T resolution a beam-spot constraint is possibly applied.

5.2 The muon spectrometer reconstruction

The Stand-Alone Muon Reconstruction starts with the estimation of the seed state from DT, CSC and RPC reconstructed segment/hits (Sec.5.2.1). Then the track is extended using an iterative method which updates the trajectory parameters at each steps and once the hits are fitted and the fake trajectories removed, the remaining tracks are extrapolated to the point of closest approach to the beam line, as explained in Sec.5.2.2.

The iterative procedure is performed using a generic interface which can be configurated in order to exclude the measurements from one or more muon subsystem. Moreover the structure of the code is such that allows the *tracker* and the *muon* code to use the same Kalman Filter algorithm and the same track parametrisation.

5.2.1 The seed generator

The algorithm first searches a pattern of DT and CSC segments using a rough geometrical criteria, then the p_T of the seed candidate is estimated using parametrisations of the form:

$$p_T = A - \frac{B}{\Delta\phi} \tag{5.3}$$

For DT seed candidates with segments in MB1 or MB2, $\Delta\phi$ is the bending angle of the segment with respect to the vertex direction. This part of the algorithm assumes the muon has been produced at the interaction point. If segments from both MB1 and MB2 exist, the weighted mean of the estimated p_T's is taken. If the seed candidate only has segments in MB3

and MB4 the difference in bending angle between the segments in the two stations is used to calculate p_T. In the CSC and overlap region, the seed candidates are built with a pair of segments in either the first and second stations or the first and third stations. $\Delta\phi$ is the difference in ϕ position between the two segments. Otherwise, the direction of the highest quality segment is used.

For the reconstruction of Cosmics, a dedicated code for the seed building has been implemented to deal with non pointing muons [37].

5.2.2 Pattern Recognition and Track Fitting

In the standard configuration the seed trajectory state parameters are propagated to the innermost compatible muon detector layer and a pre-filter is applied in the inside-out direction. Its main purpose is to refine the seed state before the true filter. The final filter in the outside-in direction is then applied and the trajectory built. The algorithm is flexible enough to perform the reconstruction starting from the outermost layer instead of the innermost.

The pre-filter and filter are based on the same iterative algorithm used in two different configurations. In both cases it can be subdivided into different substeps: search of the next compatible layer and propagation of the track parameters to it, best measurement finding and possibly update of the trajectory parameters with the information from the measurement. The process stops when the outermost (for the pre-filter) or the innermost (for the filter) compatible layer of muon detectors is reached.

At each step the track parameters are propagated from one layer of muon detectors to the next. A suitable propagator must precisely take into account material effects like multiple scattering and energy losses due to ionisation and bremsstrahlung in the muon chambers and in the return yoke. Moreover in order to reduce the processing time, the propagator must be fast. The trajectory is extrapolated in sequential steps using helix parametrisations. The required precision is obtained by using smaller steps in regions with larger magnetic field inhomogeneities. Multiple scattering and energy losses in each step are estimated from fast parametrisations,

avoiding time-consuming accesses to the detailed material and geometry descriptions. The resulting propagated state contains these effects in its parameters and errors. A number of modifications to the standard code for muons from $p - p$ collisions, described in details here [37], were necessary to take into account of these corrections for cosmic muons coming from outside the detector, in particular those traversing the detector far from its center.

The best measurement is searched for on a χ^2 basis. The χ^2 compatibility is examined at the segment level, estimating the incremental χ^2 given by the inclusion in the fit of the track segment. In case of no matching, a further selection, based again on a χ^2 threshold, is applied to determine if the hits (or segments) may anyhow be associated to the track but without including them into the fit. Then the search continues in the next station.

For the update of the trajectory parameters the pre-filter and the filter follow two different approaches. As the pre-filter should give only a first estimate of the track parameters, it uses the segment for the fit. The parameters are almost always updated as the χ^2 cut imposed at this stage is loose. The final filter instead uses the hits composing the segment with a tighter χ^2 cut which can reject individual hits. This results in a more refined trajectory state.

After the fake tracks are suppressed, the muon parameters are extrapolated to the point of closest approach to the beam line. In order to improve the momentum resolution a constraint to the beam spot (BS) is imposed, except for cosmic muons. The matrix error of the BS is diagonal and its values are: (1 mm, 1 mm, 5.3 cm).

Since the standard algorithm is applicable to a subset of cosmic muons crossing the detector close to the nominal interaction point, both dedicated cosmic-muon (CosmicSTA) and standard (ppSTA) algorithms were used to reconstruct standalone muons in CRAFT.

5.3 Muon Reconstruction with the Full CMS Detector

To ensure efficient and accurate muon reconstruction the CMS detector uses several different types of sub-detectors with different and complementary capabilities. While each sub-detector is able to measure a part of a muon's properties, the concept of a global muon is to combine information from multiple sub-detectors in order to obtain a more accurate description of the muon. In the following, the track reconstruction in the *tracker* (Sec.5.3.1) and the matching between the *tracker* track and the muon system track (Sec.5.3.2) are described.

5.3.1 Track Reconstruction in the Tracker

As in the muon system, the reconstruction process starts with the seed finding, but while in the muon system the final trajectory is built during the pattern recognition, in the tracker the pattern recognition and the final fit are performed separately.

The algorithm, the so-called *Combinatorial Kalman Filter* (CKF) [24], uses two or three consecutive hits in a defined tracker region to find the seeds. Based on the Kalman filter technique, it uses an iterative process to pass from one layer to the next and to perform the pattern recognition with a principle very similar to that used in the muon spectrometer alone (Sec.5.2). This procedure is done five times, in order to firstly search for prompt particles and then for the long-living ones.

Moreover a second algorithm, the so-called *Road Search*, which uses only the silicon strip detector to find the seeds, have been developed. It takes one hit in the inner layer and one in the outer and considers the possible paths which can connect the two initial hits. The pattern recognition is performed collecting the measurements around the paths.

5.3.2 The Global Muon Reconstruction

The track in the muon spectrometer is used to define a region of interest (ROI) in the tracker. The determination of the ROI is based on the stand-alone muon with the assumption that the muon originates from the interaction point. The definition of the region of interest has a strong impact on the reconstruction efficiency, fake rate, and CPU reconstruction time.

The ROI can be used accordly to two strategies: first, track matching with previously reconstructed tracker tracks; and second, prompt reconstruction and matching of the tracker tracks with the track in the muon system.

The latter approach has been designed for the High Level Trigger because it allows regional track reconstruction. Three reconstruction algorithms, differing in how the seeding step is performed inside the region of interest, are currently available: Inside-Out Hit-based (*IOHit*), Outside-In Hit-based (*OIHit*), and Outside-In State-based (itshape OIState). In the IOHit algorithm, the initial trajectory seed is constructed from the innermost tracker (pixel) hits by proceeding from the interaction point towards the outer edge of the silicon tracker. Unlike the IOHit strategy, the OIHit algorithm forms the trajectory seed from the outermost tracker hits, and proceeds to collect hits for the fit from the outer edge of the tracker towards the center of the detector. The OIState algorithm does not use tracker hits for seeding: it builds the trajectory seed from the extrapolation to the tracker outermost layer of a muon trajectory previously computed using only information from the muon spectrometer.

After the final global fit is made for all stand-alone track matches in the event, fake tracks are suppressed. The reconstruction of the muons ends with the matching of the global muon track and energy deposits in the calorimeters.

5.4 Muon Identification

The muon track reconstruction starts from the spectrometer and combines stand-alone muon tracks with tracks reconstructed in the inner tracker. This approach naturally identifies the muon tracks in the detector. However, a large fraction of muons with transverse momentum below 6-7 GeV/c does not leave enough hits in the muon spectrometer to be reconstructed as stand-alone muons. Moreover, some muons can escape in the gap between the wheels (Fig.2.10). A complementary approach, which starts from the tracker tracks, has therefore been designed [24] to identify off-line these muons and hence improve the muon reconstruction efficiency.

The algorithm for the muon identification of the tracker tracks extrapolates each reconstructed silicon track outward to its most probable location within each detector of interest (ECAL, HCAL, HO, muon system). After collecting the associated signals from each detector, the algorithm determines compatibility variables corresponding to how well the observed signals fit with the hypothesis that the track in the silicon detector is produced by a muon.

5.5 Muon Reconstruction Performance with Data from CRAFT

During the CRAFT data taking, dedicated reconstruction algorithms and muon identification parameters have been used [37].

At the tracker level, three different reconstruction algorithms are available. Two of them, the CKF and the *Road Search*, have been already implemented for $p-p$ collisions (see Sec.5.3.1). Moreover the *Combinatorial Kalman Filter* algorithm was configured to run in two ways: to reconstruct muons as single tracks or as two separate tracks in the two hemispheres of the detector. The third pattern recognition method, the *Cosmics Kalman Filter* (CosmicKF) has been designed to reconstruct cosmic muons crossing the tracker as single tracks. It uses a simplified Kalman Filter like trajectory building and it is able to reconstruct no more than one track per event.

All the algorithms described above end with a final fit of the collected hits, followed by the suppression of fake tracks. Moreover tracks produced by the cosmic dedicated algorithm can be split at the point of their closest approach to the beam line into two separate track candidates, with each of the candidates refit individually to yield a pair of so-called "split tracks".

Given this assortment of algorithms available to reconstruct stand-alone muon and tracker tracks, various types of global muons can be produced. The following ones have been used in the performance studies:

- *collision-like muons*, formed from 2-leg CKF tracker tracks and pp-STA standalone muon tracks. These are tracks found by the standard algorithm aimed at reconstructing muons produced in $p-p$ collisions. Only cosmic muons crossing the detector within a few centimeters of the nominal interaction point are attempted to be reconstructed by this algorithm.

- *1-leg muons*, formed from 1-leg CKF tracker tracks and cosmicSTA standalone-muon tracks. These muons typically consist of a single track in the entire tracker sandwiched between two standalone muon tracks, and yield the best estimate of the parameters of the muon.

- *split muons*, each formed from a split tracker track and a cosmicSTA standalone muon track. A cosmic muon traversing the core of the detector typically yields a pair of split muons. Comparison of these tracks, fitted independently, provides a measure of muon reconstruction performance, while the splitting mechanism ensures that they indeed comprise the same muon trajectory.

- *2-leg muons*, each formed from a 2-leg CKF tracker track and a cosmicSTAstandalone muon track. Since the two 2-leg muon tracks typically found for each tracker-pointing cosmic muon are treated independently at all stages of reconstruction, they provide fully unbiased measurements of reconstruction performance, though care must be taken to ensure that they were produced by the same muon.

An example of an event display of a cosmic muon crossing CMS (Fig.5.1) illustrates the main topological differences between tracker tracks, standalone muons, 1-leg global muons, and other types of global muons (LHC-like, split, and 2-leg).

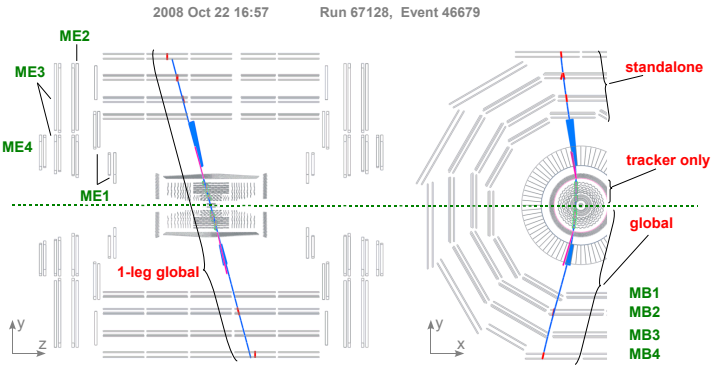


Figure 5.1: *Event display of a cosmic muon crossing CMS: the side view (left) and a part of the transverse view (right). "MB" and "ME" labels indicate positions of the muon barrel and the muon endcap stations, respectively. The solid blue curve represents a 1-leg global muon reconstructed using silicon tracker and muon system hits in the whole detector. Small green circles indicate hits in the silicon tracker. Short red stubs correspond to fitted track segments in the muon system. Energy deposits in the electromagnetic and hadron calorimeters are shown as (thin) magenta and (thick) blue bars, respectively.*

Moreover some alternative approaches were studied with CRAFT data in order to take into account multiple scattering and/or showers, produced by the muon travelling through the iron yoke, which can corrupt the measurement of its trajectory. Since using all available hits in the track fit - the approach chosen for global muons - is not always the best choice, one possibility can be for example refitting the global muon track ignoring hits in all muon stations except the innermost one containing hits. This approach is called the "tracker plus the first muon station" (TPFMS) fit.

Dedicated studies improved the momentum resolution for high-p_T muons by choosing between the results of fits including hits from the muon system or from the tracker-only fit on a track-by-track basis, depending on the fit output. One of them is called "the truncated muon reconstructor" (TMR), whereby one chooses between the TPFMS and tracker-only fits using the tail probability of the χ^2 of the fits, given the number of degrees of freedom.

Finally, during CRAFT two dedicated set of muon identification requirements, compatibility-based and cut-based, were used:

- In the compatibility-based selection, two "compatibility" variables are constructed, one based on calorimeter information and the other based on information from the muon system. A tracker muon is considered to be a muon candidate if the value of a linear combination of these variables is larger than a pre-defined threshold. Two versions of the selection, with a lower (*CompatibilityLoose*) and a higher (*CompatibilityTight*) threshold, have been taken into account for performance studies.

- In the cut-based selection, cuts are applied on the number of matched muon segments and on their proximity to the extrapolated position of the tracker track. The *LastStation* method makes use of the fact that the penetration depth of muons is larger than that of hadrons by requiring that there be well-matched segments in at least two muon stations, one of them being in the outermost station. Two versions of the *LastStation* method exist, with track-to-segment proximity cuts in only x (*LastStationLoose*) or in both x and y (*LastStationTight*) projections. In a less stringent *OneStation* method, a well-matched segment can be located in any muon station. Track-to-segment matching is performed in a local (chamber) coordinate system, where local x is the best-measured coordinate (in the $r - \phi$ plane) and local z is the coordinate perpendicular to the chamber and pointing towards the center of CMS.

5.5.1 Efficiencies

The efficiency of various muon reconstruction and identification algorithms was measured with CRAFT data [38] by selecting events with a good-quality global muon reconstructed in one hemisphere of the detector (top or bottom) and examining whether there is a corresponding track in the opposite hemisphere, in a cone of $\Delta\phi < 0.3$ and $\Delta\eta < 0.3$ around the direction of the reference track. To ensure that the muon traversed the whole detector, the p_T of the reference global-muon track was required to be larger than 7 GeV. Two-leg muons with a topology similar to that expected in $p-p$ collisions were studied; such events were selected by requiring that the distance between the point of closest approach of the reference track to the nominal $p-p$ interaction point does not exceed 10 cm in R and 20 cm in Z. Moreover a subsample of these tracks, with tighter impact parameter cuts (4 cm in R and 10 cm in Z, corresponding respectively to the beam-pipe radius and $\sim 3\sigma$ collision region at the LHC start-up) has been studied as muons produced in $p-p$ collision.

Figure 5.2 shows the efficiencies to reconstruct a) 2-leg and b) collision-like global muons and their constituents as a function of the pseudorapidity of reference muons. Integrated over the barrel region of the detector ($|\eta| < 0.8$), the efficiency for 2-leg global muons produced by the dedicated cosmic-muon reconstruction algorithm (Fig. 5.2a) was found to be $(95.4 \pm 0.3)\%$. The main source of efficiency loss is an inefficiency of about 4% in the cosmic standalone muon reconstruction, mostly in the gaps between the barrel wheels. The efficiencies of the Cosmic- CKF tracker-track reconstruction and of the tracker-track to standalone-muon matching are both larger than 99%. The efficiency of the standard global muon reconstruction algorithm in the barrel region (Fig. 5.2b), evaluated on a sample of collision like cosmic muons, was measured to be $(97.1 \pm 0.6)\%$. Since the standard stand-alone muon reconstruction was found to be $(98.7 \pm 0.3)\%$, the global muon inefficiency stems mainly from non-reconstructed tracker tracks. In fact the efficiency to reconstruct global muons grows up to $(99.7 \pm 0.1)\%$ if are taken into account only events in which both the tracker track and the standalone muon track are found.

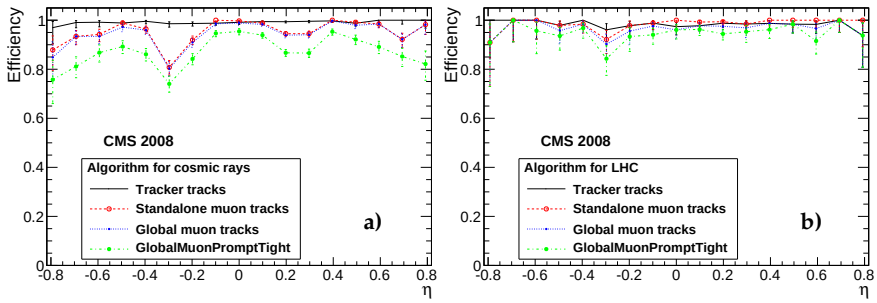


Figure 5.2: *Muon reconstruction efficiencies as a function of the ηof the reference track, for a) cosmic-muon algorithms and b) algorithms developed for muons produced in beam collisions at the LHC. Plot a) shows efficiencies of 2-leg CosmicCKF tracker tracks (solid line), Cosmic- STA standalone muons (open circles), 2-leg global muons (small filled circles), and 2-leg global muons with an additional χ^2 cut applied (large filled circles). Plot b) shows efficiencies of ppCKF tracker tracks (solid line), ppSTA standalone muons (open circles), collision-like global muons (small filled circles), and collision-like global muons with an extra χ^2cut (large filled circles).*

The efficiencies for the loose (CompatibilityLoose) and tight (CompatibilityTight) versions of the muon identification algorithm are compared with the efficiencies of tracker tracks in Fig.5.3 as a function of reference track η. For both cosmic muon and $p-p$ like track reconstruction methods, the efficiency of CompatibilityLoose tracker muons is very similar to that of the tracker tracks: the overall efficiency reduction due to CompatibilityLoose selection does not exceed 0.3%. The loss of efficiency due to the CompatibilityTight selection criteria is also small, of the order of 2%. The average efficiencies of the loose and tight versions of the LastStation variant of the cut-based selection are all above 90%.

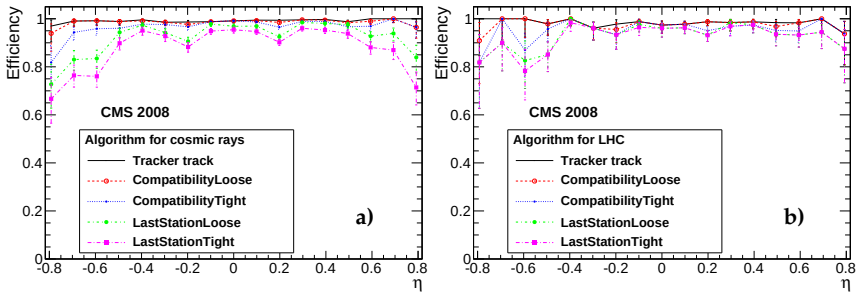


Figure 5.3: *Muon identification efficiencies as a function of h of the reference track, for a) cosmic-muon algorithms and b) algorithms developed for muons produced in beam collisions. The efficiencies for loose (CompatibilityLoose) and tight (CompatibilityTight) versions of the compatibility-based selection of tracker muons are shown in open and small filled circles, respectively. The efficiencies for loose (LastStationLoose) and tight (LastStationTight) versions of the cut-based selection of tracker muons are shown in large filled circles and in squares, respectively. For comparison, the efficiency for tracker tracks (upper line) is also shown.*

5.5.2 Momentum resolution

The muon momentum resolution was studied with CRAFT data [38] using 2-leg muons. A pure sample of muons with a topology similar to that expected in $p-p$ collisions was obtained by requiring that each of the muon tracks has at least 1 hit in the pixel detector and at least 8 hits in the silicon strip tracker. To avoid possible comparison of tracks from different muons it has been required events with exactly one pair of tracks and no hits in the CSC stations, since the alignment of the muon endcaps has not been completed

For each pair of muons the q/p_T residual, $R(q/p_T)$, is defined as

$$R(q/p_T) = \frac{(q/p_T)^{upper} - (q/p_T)^{lower}}{\sqrt{2}(q/p_T)^{lower}} \tag{5.4}$$

where $(q/p_T)^{upper}$ and $(q/p_T)^{lower}$ are the ratios of the charge sign to the transverse momentum for the upper and the lower tracks, respectively. The $\sqrt{2}$ factor accounts for the fact that the upper and lower tracks are reconstructed independently and with a similar precision. The normalized residual (or pull) for q/p_T, $P(q/p_T)$, is then defined as

$$P(q/p_T) = \frac{(q/p_T)^{upper} - (q/p_T)^{lower}}{\sqrt{\sigma^2_{(q/p_t)^{upper}} + \sigma^2_{(q/p_t)^{lower}}}} \tag{5.5}$$

where $\sigma_{(q/p_t)^{upper}}$ and $\sigma_{(q/p_t)^{lower}}$ are the estimtes of q/p_T errors for the upper and the lower muons, respectively.

The width of $P(q/p_T)$ distributions was examined to verify the accuracy of the estimates of track parameter errors. These estimates depend on, among other things, the so-called alignment position errors (APEs) accounting for the precision with which the positions of different detector components are known. The available sample of 2-leg muons was divided into several subsamples according to the p_T value of the muon reconstructed in the lower hemisphere, and a Gaussian fit to the $P(q/p_T)$ distribution for each subset was performed. Figure 5.4 shows the widths of these Gaussian fits as a function of the reference track p_T, for 2-leg muons reconstructed by various muon reconstruction algorithms outlined in Section 5.3.2. If all errors were calculated correctly, these widths should be 1.0. The widths of the pulls for standalone muons are greater than unity at all p_T values because the muon APEs, which were not yet fully implemented, were all set to zero in the reconstruction. The widths of other pulls are consistent with unity in the region of $p_T < 40$ GeV/c, confirming that the estimates of errors for the low-p_T region are accurate. In the higher-p_T region, the widths of the tracker-only pulls are larger than 1.0, indicating that the tracker APEs are underestimated. As the muon p_T increases, so does the importance of the muon system in the momentum measurement, and the widths of the pulls for the combined tracker-muon fits move closer to the widths of the pulls of the standalone-muon fit.

The same subsets of 2-leg muons have been used to evaluate the q/p_T resolution in various bins of p_T. Figure 5.5 shows the widths of the Gaussian

fits to $R(q/p_T)$ distributions obtained with various muon reconstruction algorithms. In the p_T region below approximately 200 GeV/c, where the resolution is dominated by multiple scattering effects, the inclusion of muon hits does not improve the resolution beyond that obtained with the tracker-only fits.In the high p_T region, the resolution obtained with TPFMS and TMR is better than the global muons and of tracker-only tracks ones, as expected.

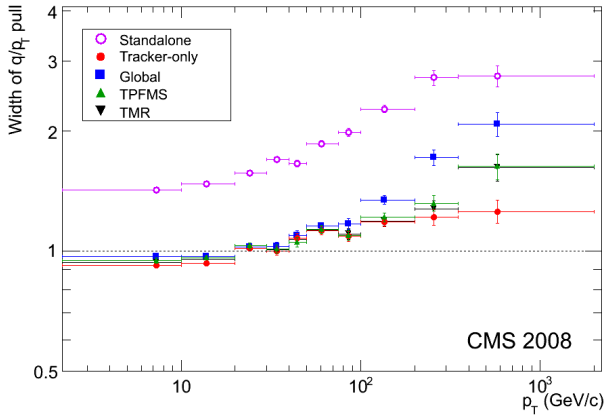


Figure 5.4: *Width of Gaussian fits to q/p_T pulls for various muon reconstruction algorithms as a function of reference track p_T.*

5.5.3 Charge assignment

The tracker and muon systems of CMS are expected to misidentify muon charge sign at a very low rate, typically below 1% of the time. Extremely high energy muons are an exception since because of their nearly straight trajectory their misidentification probability is higher.

The rate of charge misassignment with CRAFT data has been studied [38] by measuring the number of times that two measurements of the charge of the same muon, in the top and bottom hemispheres, disagree. In order to obtain two independent measurements of the muon charge, 2-leg muons

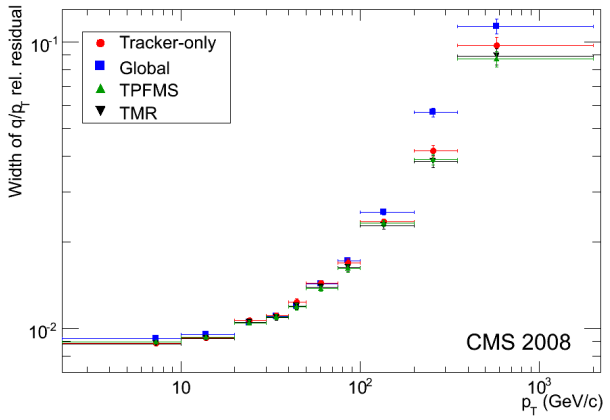


Figure 5.5: *Width of Gaussian fits to q/p_T resolutions for various muon reconstruction algorithms as a function of the reference track p_T.*

in the tracker pointing dataset have been used. Since the subject under study is a low-rate effect, the purity of the test sample is crucial. For this reason some additional cuts, below listed, have been required to the events selected for the analysis.

- Since the dominant background comes from muon showers, where multiple muons are traversing the detector at the same time and the top and bottom measurements may not correspond to the same muon, at most two cosmicSTA muons have to be present in the event. Moreover these two muons are required to be in the barrel muon spectrometer, one in the top hemisphere, the other in the bottom.

- To obtain a pure sample of barrel muons, a veto cosmic rays in the muon and tracker endcaps was imposed by requiring that there are no CSC or tracker endcap hits in the muon fit.

- In order to guarantee good fits of transverse momentum and charge, at least 5 hits in the tracker outer barrel are required.

- To select a sample of muons resembling those expected from beam collisions, the point-of-the-closest-approach of each track to the nominal beam line is required to lie within $r < 50$ cm and $|z| < 30$ cm of the nominal position of $p - p$ interactions.
- Finally, the charges of the tag track assigned by the tracker-only, global, and TPFMS algorithms must all be the same.

Figure 5.6 shows the charge misassignment ratio in bins of transverse momentum of the tracker track reconstructed in the top hemisphere for different muon reconstruction algorithms. As expected, the measurement of the charge provided by the standalone muon reconstruction is less accurate than the tracker one in the whole p_T range. Both the trackeronly fit and the combined tracker-muon fits provide a reliable charge measurement for the low momentum region. At high-p_T values, the most accurate charge assignment is given by the dedicated high-p_T muon reconstruction algorithms. Is also import to remark that while different algorithms lead in performance over different momentum regions, a charge misassignment of 1% or better is achieved over the entire studied momentum spectrum, as expected.

5.5.4 Muon identification

As discussed in Section 5.4, an approach complementary to the global muon reconstruction is to consider all tracker tracks above a certain momentum to be potential muon candidates, and to check this hypothesis by looking for compatible signatures in the calorimeters and in the muon system. A crucial step of this algorithm is the track propagation which has to take into account the magnetic field, the average expected energy losses in the detector materials, and multiple scattering.

The accuracy of the propagation and the performance of the track-to-segment match were studied using the tracker pointing muon dataset. Every track reconstructed by the CosmicTF algorithm with $p_T > 1.5$ GeV/c or $p > 3$ GeV/c was propagated to the layers of the muon chambers, and a search for the nearest muon segment reconstructed in each layer was

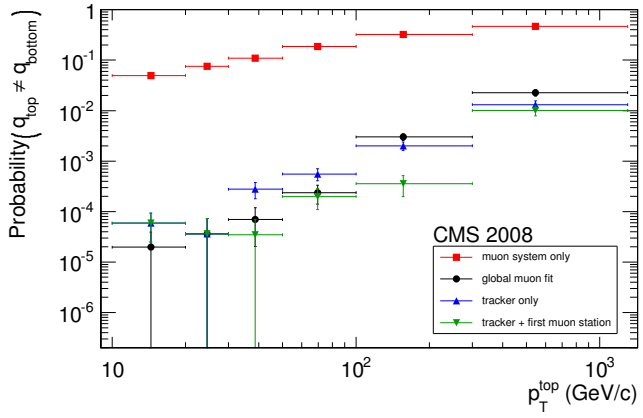


Figure 5.6: *The charge misassignment ratio as a function of p_T of the tracker track reconstructed in the top hemisphere, for standalone muons (red squares), tracker tracks (blue triangles), global muons (black circles), and the TPFMS refit (green upside-down triangles).*

performed. For each segment found, the normalized residuals (pulls) in position and in slope were calculated; the pull is defined as the difference between the position (or slope) of the extrapolated track and the position (or slope) of the nearest segment divided by their combined uncertainty. The track and the segment are considered to be matched if the distance between them in the local (*i.e.* calculated in the reference frame of the chamber) x is less than 3 cm or if the value of the pull for local x is less than 4.

Figure 5.7 shows the distributions of the distance in the x direction between the extrapolated track position and the position of the segment for successful track-to-segment matches separately in three bins of tracker track p_T: less than 20 GeV/c, 20-50 GeV/c, and above 50 GeV/c. As expected, the width of the distributions increases with the extrapolation distance, from the innermost to the outer most muon stations (from MB1 to MB4 and from ME1 to ME3 in the DT and the CSC systems, respectively) and with decreasing p_T due to the larger impact of multiple scattering effects.

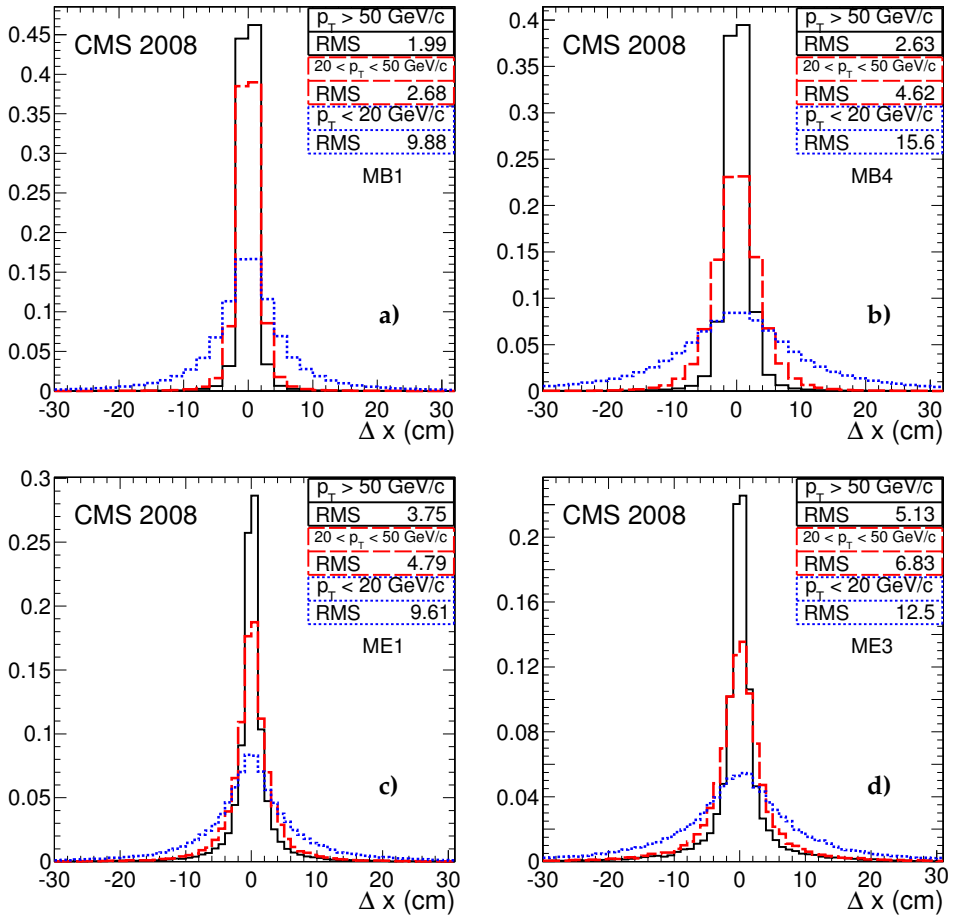


Figure 5.7: *Residuals distributions of the local x position for the track-to-segment match in CRAFT data, shown separately in three p_T bins: less than 20 GeV/c (solid blue histogram), 20-50 GeV/c (red dashed histograms), and above 50 GeV/c (dotted black histograms). Four panels show residuals in a) MB1 b) MB4 c) ME1 d) ME3 chambers. Each histogram is normalized to unit area.*

The pulls of the local x position and of the local $\mathrm{d}x/\mathrm{d}z$ slope in the DT and CSC systems are shown in Fig. 5.8. The widths are close to unity and no large biases were observed, thus demonstrating that the propagation works as expected and the uncertainties are estimated correctly.

Finally, Figure 5.9 shows the probability of a successful track-to-segment match in the DT system as a function of the distance between the propagated track position and the nearest chamber edge. The distance to the chamber edge is defined to be negative when the extrapolated position of the track is inside the nominal chamber volume, and to be positive otherwise. As expected, the efficiency of finding a muon segment well inside the chamber is close to unity. The slope of the efficiency drop at the chamber edge is consistent with that expected from the Monte Carlo studies.

5.6 The Muon Reconstruction Offline Monitoring and Certification

As a general overview of the CMS Data Quality Monitoring as been done in Section 4.3.1, here the main aims of the Offline DQM system are presented (Sec.5.6.1). Since I devoted a great part of my work in the implementation of the offline monitoring and validation Muon Reconstruction software, a detailed description of these subjects is given in Section 5.6.2. Finally, in Section 5.6.3 is described the present status and the ongoing work on the data certification process.

5.6.1 The offline DQM system

5.6.1.1 Data processing

Numerous offline workflows in CMS involve data quality monitoring: Tier-0 prompt reconstruction, re-reconstruction at the Tier-1s and the validation of the alignment and calibration results, the software releases and all the simulated data. These systems vary considerably in location, data content and timing, but as far as DQM is concerned, CMS has standardised on a single two-step process for all these activities, shown in Fig.5.10.

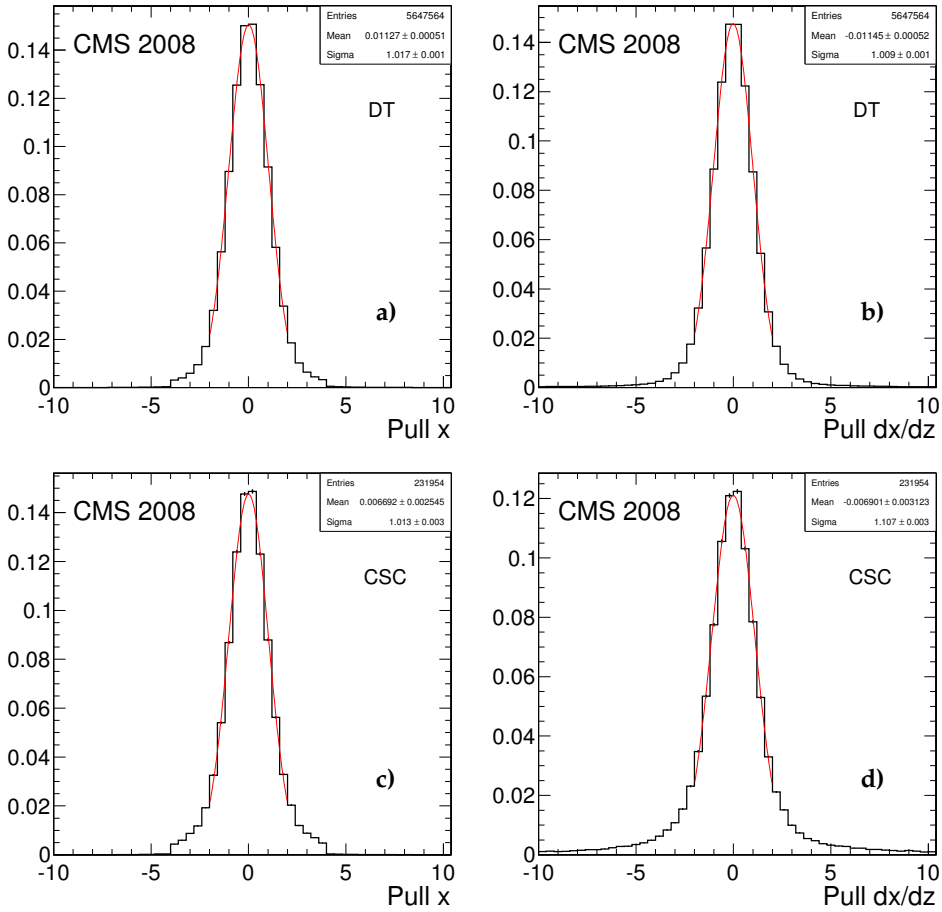


Figure 5.8: *Pulls distributions of the local x position and the local dx/dz slope of the track-to- segment match: a) pull of x in the DT system; b) pull of dx/dz in the DT system; c) pull of x in the CSC system; d) pull of dx/dz in the CSC system. Each plot is a combined distribution of pulls in all MB or ME stations. All histograms are normalized to unit area; the superimposed curves are the results of Gaussian fits in the range -2,2.*

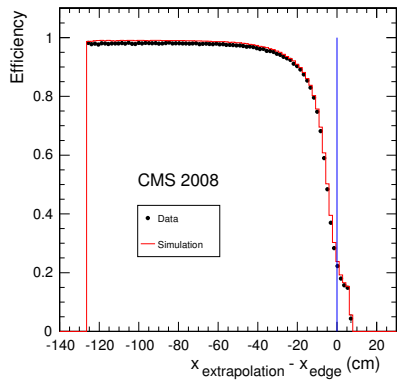


Figure 5.9: *Efficiency of a successful track-to-segment match in the DT system, averaged over all DT chambers, as a function of the distance between the propagated track position and the nearest chamber edge, for CRAFT data (points) and the Monte Carlo simulation (histogram).*

In the first step the histogram monitor elements are created and filled with information from the CMS event data. The histograms are stored as *run products* along with the processed events to the normal output event data files. When the CMS data processing systems merge these output files, the histograms are automatically summed together to form the first partial result.

In a second step (the so-called *harvesting*), which is run at least once at the end of the data processing and sometimes periodically during it, the histograms are extracted from the event data files and summed together across the entire run to yield the full event statistics. Moreover detector control system (DCS, in particular high-voltage system) and data acquisition (DAQ) status information are received from the online condition database. The application analyses all these quantities using detector specific algorithms and creates new histograms such as high-level detector or physics object summaries.

The final histograms are used to calculate efficiencies and checked for quality, in particular compared against reference distributions. The harve-

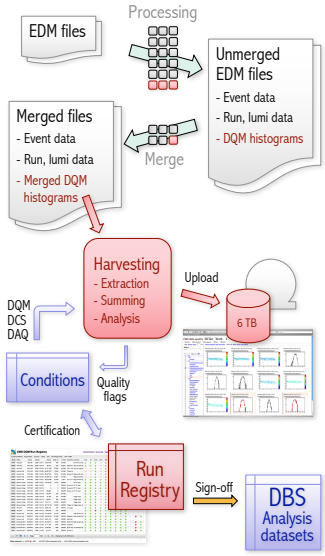


Figure 5.10: *The offline DQM system.*

sting algorithms also compute the preliminary automatic data certification decision. The histograms, certification output and quality test results along with any alarms are written into a ROOT file, which is then uploaded to a central DQM GUI web server.

5.6.1.2 Visualization

As in the online case, the DQM results from offline processing are uploaded to the central DQM GUI server with a large disk pool. There the result files are merged to larger size and backed up to the tape; recent data is kept cached on disk for several months. The automatic certification results from the harvesting, called quality flags, are extracted and uploaded to the run registry. From there the values are propagated to the condition database and the dataset bookkeeping system (DBS).

CMS provides one central DQM GUI web server instance per offline activity, including one public test instance for development. All online and offline servers provide a common look and feel and are linked together as one entity, to give the entire worldwide collaboration access to inspect and analyse all the DQM data at one central location.

5.6.1.3 Operation

In all the offline processing the initial histogram production step is incorporated in the standard data processing workflow as an additional execution sequence, using components from standard software releases. The harvesting step implementation currently varies by the activity, there are some manual operations but the final system will rely on a fully automated process including the histograms upload to the DQM GUI and the quality flag writing on the Run Registry.

5.6.2 The Muon Reconstruction Offline Monitoring

A reconstructed muon, the so-called *RecoMuon*, contains not only the track fit parameters but also many other information as the energy deposits and the muon identification variables. Moreover, as described in Sec.5.2 and 5.3, muons can be reconstructed using only RecHits from the spectrometer (StandAlone muon), the Tracker together with the identification parameters (Tracker Muon) or both the muon and tracker systems (Global muon). Fig.5.11 summarises in a graphical view all these muon types and their properties; the offline monitoring produces and automatically tests a list of histograms allowing the shifters able to understand the status of each box.

5.6.2.1 The Run Products

At present the plots available concern:

- *track specific variables*, as the track quality (in terms of χ^2, number of RecHits used and lost during the fitting procedure), the distance of the closest approach and the momentum (components and errors).

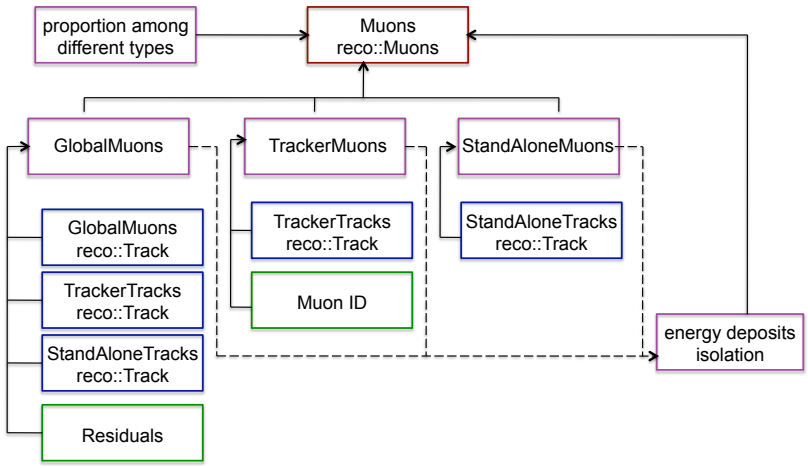


Figure 5.11: *Main components of the three reconstructed muon types.*

Moreover the correlations between these parameters and the pseudorapidity, the azimuthal direction and inverse transverse momentum of the reconstructed muon are monitored.

- *comparison among the different parts which compose the global muon*: residuals on charge, pseudorapidity, inverse transverse momentum, inverse momentum, radial and azimuthal direction are computed between all the pairs of muon-alone, tracker-alone and global tracks.

- *track seeds*. The seed direction, number of RecHits and momentum are plotted together with their correlations with the associated seed pseudorapidity, azimuthal direction and inverse transverse momentum.

- *recHits*. Their provenance is checked with different level of granularity as the percentage of hits not used for the track fit respect to the total number of hits available in the event.

- *energy deposits*: distributions are made separately for the electromagnetic/ hadronic calorimeter, for the barrel/ endcap regions and for single cells and 3×3 towers. These plots are dominated by the background noise in comics data samples because muons are not in general pointing to the vertex crossing only a small region of the calorimeters and leaving small signals. For this reason some dedicated plots, made only with pointing muons, have been developed for the cosmic commissioning period.

- *muon identification quantities*, as the number of chambers with matching segments divided the number of total track associated segments (also separately for DTs and CSCs), the residuals and pulls (and their errors) between the segments used for the track fit and the track itself.

In Figure 5.12 there are some example plots taken from a CRAFT cosmic run and referred to global muons properties. The first distribution on the top-left side shows the muon pseudorapidity distribution; the drop on the left part is due to the presence of the CMS cavern shaft which deflects the trajectories of muons crossing it. The muons track quality, in terms of $\chi^2/d.o.f.$, is represented in the second plot on the right side; the maximum value of the distribution is clearly around 1. In the second line there are the residuals on pseudorapidity between the stand-alone (left) and the tracker (right) tracks which are part of the global track and the global track itself; the distributions are symmetric around 0 and the best width is clearly associated to the tracker track, as expected. In the following line there are two plots related to the hits (left) and segments (right) provenance; due to the large amount of muons passing through the barrel, as respect to the endcaps, most of the tracks contain only hits from DTs and RPCs and only segments coming from four DTs stations. In the bottom part of Fig.5.12 two plots show the energy deposits of pointing muons in the 3×3 towers of ECAL (note the peak around 300MeV, as expected from a *Minimum Ionizating Particle* crossing the electromagnetic calorimeter) and the number of muon chamber matching which the global track, variable used for muon identification studies.

5.6.2.2 The harvesting step

In the second step of the offline data processing the most relevant histograms are automatically tested. The results are float numbers from 0 to 1 stored in summary plots. This allows a quick evaluation of the data quality, also for non-expert shifters.

The muon reconstrction actually produces 5 summary histograms. They concern :

- *Kinematics parameters.* Three distributions are checked: the χ^2 reduced is required to have its maximum inside a defined range; the pseudorapidity has to be symmetric respect to 0 (taking into account the shaft effect for cosmic data); the radial direction is checked to be negative for Coscmics and symmetric respect to 0 in the case of $p-p$ collisions. These tests are done for all the 5 type of tracks available: the stand-alone and the tracker ones associated to the global muon, the global track and the stand-alone and the tracker ones.

- *Residuals among the different global muon tracks.* The parameters tested are: pseudorapidity, radial direction and inverse transverse momentum. Residuals are computed for all the possible combinations: stand-alone versus global tracks, tracker versus global tracks and finally stand-alone versus tracker tracks. The width of these distributions is required to be inside a defined range. Moreover the charge difference among the three types of tracks is checked and the percentage of non-agreements is required to be less than a certain threshold.

- *Muon identification parameters.* For both tracker and global muons three different distributions are tested: the maximum number of matching segments must be 4; the percentage of matching segments coming from one single detector respect to the total number of track associated segments have to be over a certain threshold, which is in general different for DT and CSC stations; the width of the normalized residuals (pulls) between segments and track is required to be around 1 for both the x and y projections.

- *Energy deposits.* The energy deposit distributions (in the 3×3 towers) for both the electromagnetic and the hadronic calorimeters are fitted with the convolution of a gaussian and a landau functions and then it is checked their most probable value (MPV). This value is expected to be around 240 MeV for ECAL and 2.5 GeV for HCAL. To avoid the huge noise background, only the sub-sample of pointing muons is used with cosmic data. The test is applied separately on all the three types of muons: global, stand-alone and tracker ones.

- *Muon moltelpicity.* A final check is done on the percentage (respect to the total) of each type of reconstructed muon, in order to understand if there are one or more problematic subsystems which affects the fitting procedure.

Studies to tune the tests thresholds and validity ranges for both cosmic and proton-proton collision data are still on going. They are based on the comparison with simulated data and detector performance analysis. Moreover the daily shift reports during the cosmic commissioning have already played a key role for the quality tests improvement and they will be crucially important during the LHC start-up period.

5.6.3 The Muon Reconstruction Certification

CMS uses a run registry database with a front-end web application as the central workflow tracking and bookkeeping tool to manage the creation of the final physics dataset certification result. The certification process begins with the physicists on online and offline shift filling in the run registry with basic run information, and adding any pertinent observations on the run during the shift. This information is then augmented with the automatic data certification results from the online, Tier-0 and Tier-1 data processing as described in the previous sections. For each subdetector and physics high-level objects (as muons, electrons, jets ...) one single boolean flag or a floating point value describes the final quality result. For the latter report, appropriate thresholds which yield binary “good” or “bad” results are applied. The result “unknown” is labelled if no quality flag was calculated.

Once the automatic certification results are known and uploaded to the run registry, the person on shift evaluates the detector and physics object quality following the shift instructions on histograms specifically tailored to catch relevant problems. This person adds any final observations to the run registry and overrides the automatic result with a manual certification decision where necessary.

The best metrics to merge all the information in few flags containing the data quality of so many different physics objects is at present under study. For what concern the muon reconstruction the main open question is actually how to assign a weight, which could represent a "degree of importance", to the results of the tests described in the previous section for the final quality flag computation.

A great effort coming from people involved in code improvements, shifts and offline analysis is bringing the full Data Quality Monitoring to a robust and reliable system, at least within the LHC start-up period.

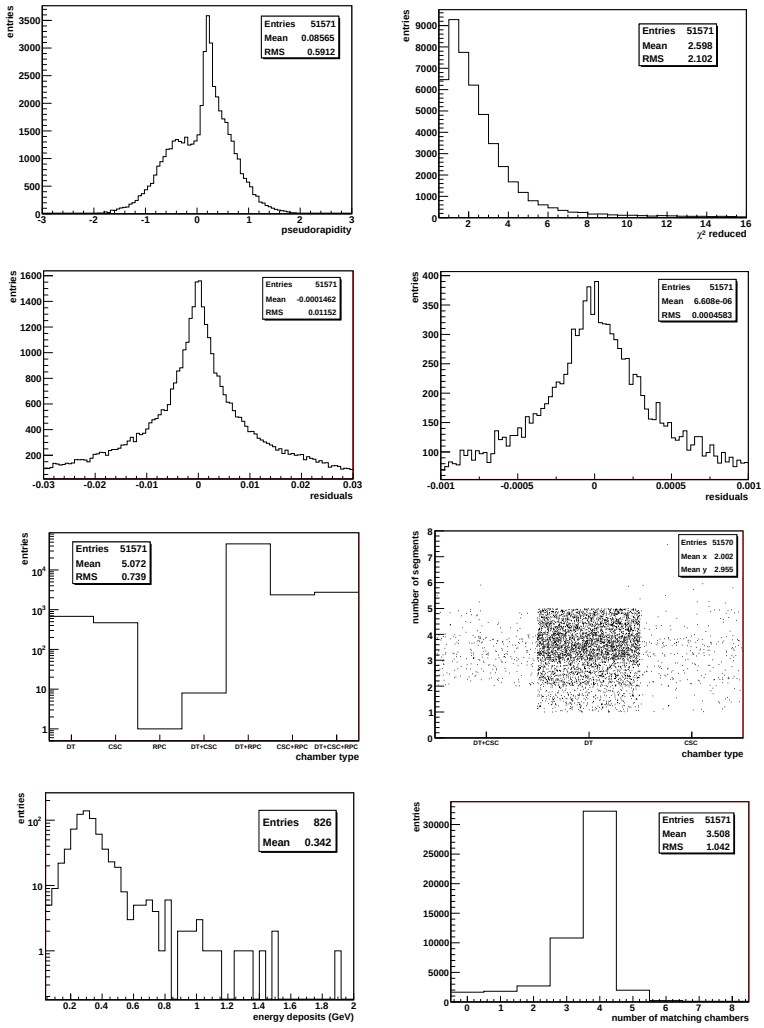


Figure 5.12: *Histograms concerning global muon properties taken from the DQM offline used during the CRAFT period. From the top-left to the bottom-right: pseudorapidity and chi^2 reduced distribution; residuals between stand-alone and tracker global tracks and the global track itself; hit and segments provenance from the muon system; ECAL energy deposits of pointing muons and number of muon chamber with segments matching with the global track.*

Chapter 6

The Higgs search through the H→ZZ→ 4μ channel

The inclusive production of the Higgs boson followed by the decay into four leptons (H→ZZ→ $l^{\pm}l^{\mp}l'^{\pm}l'^{\mp}$, with $l, l' = e$ or μ) is expected to be one of the main discovery channels at the LHC over a wide range of possible m_H values.

In the present chapter an analysis strategy is presented to spotlight the discovery potential of the 4μ final state in the context of the startup luminosity and centre of mass energy at the LHC. This study is part of a more general analysis on the H→ZZ→ $4l$ signature, described in details in Ref.[39] for a 14 TeV centre of mass energy. The aim of my work was to check the previous analysis with the total amount of data that will be collected at the LHC during the start-up period with the $p - p$ interaction energy of 10 TeV. For completeness, the event MC production and the preliminary event selection requirements, common for all the three possibile lepton decay modes, are presented in the context of the complete H→ZZ→ 4ℓ analysis strategy. The main results obtained combining the three $4l$ channels are shown too. Emphasis is put on the reduction of distinguishable background rates and on methods allowing a data-driven derivation of experimental and background systematic uncertainties. This

analysis investigates the CMS sensitivity for the observation of a Higgs boson with mass in the range from 115 GeV/c^2 to 250 GeV/c^2, and it is performed basically using a sequential set of cuts.

6.1 Monte Carlo simulation of the relevant physics processes

A complete $4l$ final state production of signal and background datasets obtained with a detailed Monte Carlo simulation [40] of the detector response have been preliminary done. All the samples have been then subject to full reconstruction [41].

The general multi-purpose Monte Carlo event generator PYTHIA [3] have been used for the various signal and background processes described in details in the following, either to generate a given hard process at leading order (LO), or only for the showering and hadronization in cases where other Monte Carlo have been used for the primary event production. All signal and background processes have been re-weigthed to next-to-leading-order (NLO) cross sections.

The main processes considered for the $4l$ analysis are reported in Table 6.1. Note that here and henceforward, Z stands for Z, Z*, and γ^* (where possible). Additional details on the production of the corresponding samples are given in sections 6.1.1 to 6.1.4. Moreover other selected control samples have been used to derive from data a normalization for the main backgrounds, and to control efficiencies and systematics.

6.1.1 Signal: $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$

The Higgs boson samples have been generated with PYTHIA (LO gluon and weak-boson fusion, $gg \rightarrow H$ and $q\bar{q} \rightarrow q\bar{q}H$). The parton density function (PDF) set CTEQ5L with the QCD scale set at PYTHIA's default values is used. The Higgs boson is forced to decay to two Z-bosons, which are allowed to be off-shell, and both Z-bosons are forced to decay via $Z \rightarrow 2\ell$.

Events have been then re-weighted to the total NLO cross-section

Process	MC	$\sigma_{NLO}\, BR$
$H \to ZZ \to 4\ell$	PYTHIA	2-25 fb
$t\bar{t} \to 2Wb\bar{b}$	MadGraph	281 pb
$Zb\bar{b} \to 4\ell$	MadGraph	56 pb
$ZZ \to 4\ell$	MadGraph	0.189 pb

Table 6.1: *Monte Carlo simulation datasets used for cut optimization;* Z *stands for* Z, Z*, γ^*; ℓ *means* e, μ, τ.

$\sigma_{NLO}(pp \to H) \cdot BR(H \to ZZ^{(*)}) \cdot BR(Z \to 2\ell)^2$, where $\sigma_{NLO}(pp \to H)$ and $BR(H \to ZZ^{(*)})$ were taken from [11, 42] and $BR(Z \to 2\ell) = 0.101$ [43]. Figure 6.1 shows this $H \to 4\ell$ cross-section as a function of the Higgs boson mass m_H. In this analysis a set of 20 Monte Carlo samples corresponding to different Higgs boson masses: from 115-205 GeV/c^2 with a step of 5 GeV/c^2, with 1 additional point up to 250 GeV/c^2, has been used.

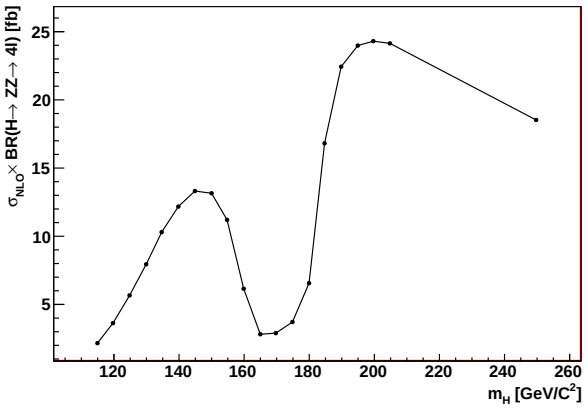


Figure 6.1: *The NLO cross-section for* $H \to 4\ell$ *as a function of the mass* m_H *calculated as* $\sigma_{NLO}(pp \to H) \times BR(H \to ZZ^{(*)}) \times BR(Z \to 2\ell)^2$, *where* ℓ *stands for* e *and* μ.

6.1.2 Background: $t\bar{t}$

A full inclusive sample $t\bar{t}$ + $n$jets ($n = 0, 1, 2, 3$) has been generated with MadGraph [44], using m_{top}=170.9 GeV/c^2 and the PDF set CTEQ6L1. Parton showering and hadronization has been done by PYTHIA. Events have been preselected with the following requirements: at least four leptons (electrons and/or muons) with $p_T > 4$ GeV/c within $|\eta| < 2.7$. The efficiency of this preselection on the inclusive sample is 0.01091. The LO cross section is 280.9 pb and the statistics survived to the filter corresponds to a integrated luminosity of 3.6 fb^{-1}. The NLO K factor is 1.46, further details on its computation can be found here [45].

6.1.3 Background: $Zb\bar{b} \to 4\ell$

The $Zb\bar{b} \to 4\ell$ sample has been generated with MadGraph event generator, interfaced to PYTHIA for showering and hadronization. The corresponding LO cross-section is 56 pb. To obtain the NLO cross-section, a K-factor using MCFM [46] has been calculated: $K_{NLO} = 1.66$. In this analysis, a preselected sample is used with the same requirements as for the $t\bar{t}$ background. The preselection efficiency is resulted 0.007. The $Zb\bar{b}$ simulated sample corresponds to an integrated luminosity of about 19 fb^{-1}.

6.1.4 Background: $q\bar{q} \to ZZ^{(*)} \to 4\ell$

The $q\bar{q} \to ZZ^{(*)} \to 4\ell$ sample has been generated with MadGraph, then it has been further interfaced to PYTHIA for showering and hadronization. The MadGraph LO cross-section is 189 fb. As for the background samples described above, the requirement of the presence of four leptons with p_T >4 GeV/c inside the detector acceptance has been applied, with a resulting efficiency of 0.3165. The final $ZZ^{(*)}$ sample corresponds to an integrated luminosity of about 4.6 pb^{-1}. To account for contributions to all the NLO diagrams and the NNLO gluon fusion process ($gg \to ZZ$, known to contribute $\approx$ 20% with respect to the LO [47]), events have been reweighted with a m_{4l}-dependent K-factor $K(m_{4\ell}) = K_{NLO}(m_{4\ell}) + 0.2$. The NLO K-factor $K_{NLO}(m_{4\ell})$, obtained with MCFM [46], corresponds to an

average correction of $< K >= 1.35 + 0.2 = 1.55$ as shown in Fig.6.2. Details on the calculation of this $m_{4\ell}$-dependent K-factor and other dynamic differences between NLO and LO processes are presented here [48].

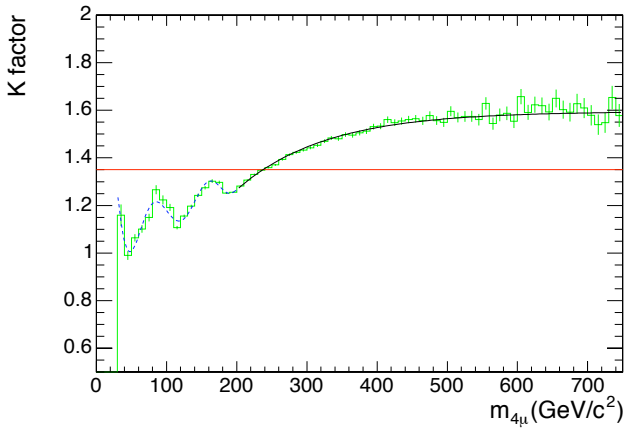


Figure 6.2: *Mass-dependent Next-to-Leading-Order K-factor $K_{NLO}(m_{4\ell})$ for the $ZZ^{(*)} \rightarrow 4\ell$ process as evaluated with MCFM [46].*

6.2 Trigger selection

For Higgs boson masses m_H above 100 GeV/c^2, the intermediate state in the cascade $H \rightarrow ZZ^{\star} \rightarrow 4\ell$ is expected to be dominantly produced with at least one Z boson on the mass shell, which then decay in pair of leptons carrying each a p_T of about $m_Z/2$. The triggering of the CMS detector on the Higgs boson signal relies on the presence of one or two high p_T leptons.

For the LHC start-up luminosity of $\mathcal{L} = 10^{32}$ cm^{-2}s^{-1}, the High Level Trigger (HLT) configuration foreseen in CMS allows for single lepton p_T thresholds well below 20 GeV/c. Hence a very high selection efficiency is expected for the Higgs boson.

The E_T thresholds for the different lepton trigger paths are reported in Table 6.2. A global "OR" between different HLT sequences (*trigger-paths*)

HLT path	HLT threshold [GeV]
Single Isolated e	15
Single Relaxed e	18
Double Isolated e	(10,10)
Double Relaxed e	(12,12)
Single Isolated μ	11
Single Relaxed μ	16
Double Relaxed μ	(3,3)
Double Isolated $e, , \mu$	(8 , 7)
Double Relaxed e, μ	(10 , 10)

Table 6.2: *E_T thresholds for the different lepton trigger paths at the High Level Trigger (HLT). The "relaxed" lepton triggers imply loose isolation in the case of electron and no isolation requirements in the case or muons.*

is chosen to maximize the signal detection efficiency. Processes including electrons are considered also for the 4 muons channel, in order to have a common trigger selction for all the three 4ℓ channels. Therefore the triggers-paths taken into consideration are: single muon isolated, single muon (no isolation), double muon (no isolation), single electron isolated, single electron (relaxed isolation), double electron isolated, double electron (relaxed isolation) and their combinations. The double muon isolated path is not used as it is essentially redundant with the corresponding non isolated path for what concerns signal selection.

The HLT path efficiency for signal and its error are defined as $\varepsilon = (\#\ L1 \wedge HLT\ evt)/N_{gen}$ and $\delta\varepsilon = \sqrt{(\varepsilon(1-\varepsilon))/(N_{gen})}$, where N_{gen} is the number of generated events in the sample having four leptons of the right charge ($2\mu^-2\mu^+$, $2e^-2e^+$ or $\mu^-\mu^+e^-e^+$) within the detector acceptance, i.e. $|\eta(\ell)| < 2.5$. The trigger efficiencies for the main backgrounds are evaluated starting from samples produced with the following requirements at generation level:

- ZZ: at least 4 leptons (muons or electrons);

- $Zb\bar{b}$: at least 4 leptons (muons or electrons) with $p_T > 2$ GeV/c and $|\eta| < 2.7$ (filter efficiency: 1.6%);
- $t\bar{t}$: at least 4 leptons (muons or electrons) with $p_T > 2$ GeV/c and $|\eta| < 2.7$ (filter efficiency: 8.6%).

The trigger efficiencies are reported in Fig. 6.3 for the $H \rightarrow 4\mu$ channel, compared to the $H \rightarrow 4e$ and $H \rightarrow 2\mu 2e$ channels, and for several hypothetical values of the SM Higgs boson mass, as well as for the main backgrounds, ZZ, $Zb\bar{b}$ and $t\bar{t}$. A very high trigger efficiency is obtained for the signal, with values above 99% for a Higgs boson mass m_{H} above 115 GeV/c^2 and reaching close to 100% for m_{H} above 200 GeV/c^2. The rate of main background events which pass the HLT selection is estimated to be of about 3998 events/fb^{-1} of $t\bar{t}$, 597 events/fb^{-1} of $Zb\bar{b}$ and 90 events/fb^{-1} of ZZ.

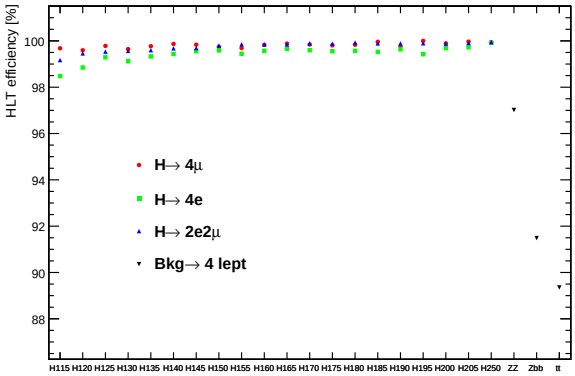


Figure 6.3: *Expected trigger efficiencies for the three $H \rightarrow 4\ell$ channels and the main backgrounds for a logical OR of all the single and double lepton trigger paths of the CMS High Level Trigger at LHC start-up luminosities.*

6.3 Lepton properties

The reconstruction of the SM Higgs boson in the decay chain $H \rightarrow \mathrm{ZZ}^* \rightarrow 4\ell$ imposes high performance reconstruction, identification and isolation as well as excellent energy-momentum measurements of leptons. The identification of isolated leptons emerging from the event primary vertex allows for a drastic reduction of QCD-induced sources of misidentified ("fake") leptons. Therefore precision energy-momentum measurements translate directly in an accurate precision Higgs boson mass measurement, which is the single most discriminating observable for the Higgs boson search.

With four leptons in the final state, and in view of the modest fraction of the total production cross-section observable in the 4ℓ channels, a very high lepton reconstruction efficiency is mandatory. This turns out to be especially challenging for the reconstruction of leptons at very low p_T^ℓ. For Higgs bosons with masses $m_\mathrm{H} \lesssim 2m_Z$, one lepton pair at least couples to an off-shell Z^* boson. The softest lepton in that pair typically has $p_T^\ell \sim 10$ GeV/c for masses $m_\mathrm{H} < 140$ GeV/c^2. Such very low p_T^ℓ values lie at the extreme edge of the domain which will be controlled at the LHC using tag-and-probe methods in inclusive single m_Z production. In the low p_T range, a full combination of information provided by the tracker and electromagnetic calorimetry (for electrons) or by the tracker and muon spectrometer (for muons) becomes essential for the reconstruction, identification and isolation of leptons.

Since a detailed description of the muon reconstruction and identification performance has been already done in Chapter 5 and the main performance results with Monte Carlo data can be found here [49], the sections 6.3.1 and 6.3.2 focus on the muon isolation and its measurement uncertainties. The comprehension of these three variables has been crucial to define the analysis kinematics cuts described in the following for the 4μ channel. Similar studies have been done for electrons. Further details can be found in the reference of the complete 4ℓ analysis note [39].

6.3.1 Muon Isolation parameters

To compute the muon isolation parameter an isolation cone is defined in the $\eta - \phi$ space around each reconstructed muon. The cone radius was chosed $R_{iso} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. An inner 'veto' cone is also defined around the muon, in order to subtract the track (and energy deposits) of the muon itself from the overall amount. The sizes of both the isolation and the inner veto cones have been optimized for this analysis and the values $R_{\rm iso} = 0.3$ and $R_{\rm veto} = 0.015$ are found to provide the best working point for the analysis event selection.

If two or more muons fall in the same isolation cone, the contributions of the extra muon(s) is subtracted. Similarly, the effects of muon bremsstrahlung are properly taken into account and corrections are applied.

Thresholds on the p_T, on the radial and longitudinal impact parameters of the tracks and on the corresponding significancies have also been optimized with the goal to suppress ghost tracks. They depend on the number of hits per track: the more track hits, the more a track is likely to be a real one. Impact parameter selections are very effective especially against $\mathrm{Z}b\bar{b}$ background. The efficiency of these cuts is estimated to be at the level of 98%. The sum of the deposits inside an isolation cone runs only over the tracks that survive these preliminary cuts.

The calorimeter-based observables refer to the ECAL ($\mu\mathrm{Iso}_{\mathrm{ECAL}}$) and/or to HCAL ($\mu\mathrm{Iso}_{\mathrm{HCAL}}$) deposits in the cone around the muon track. Both track based ($\mu\mathrm{Iso}_{\mathrm{track}}$) and calorimeter based isolations are combined in a single isolation variable for the final muon selection (Sec 6.5.1).

6.3.2 Muon measurement uncertainties

Uncertainties from reconstruction

A data-driven method for the estimation of systematic uncertainties on the muon reconstruction efficiency relies on the usage of a sample of inclusive W and Z bosons for a single "tag" muon with the requirement of $p_T > 19$ GeV/c. By counting the number of $Z \to 2\mu$ events in the reso-

nance peak of the invariant mass distributions built using the tag muon and either all other tracks, or all other standalone muons, or all other globally reconstructed muons, one can evaluate the efficiency of finding globally-reconstructed muons. The efficiency of finding globally-reconstructed muons can thus be measured with an uncertainty of better than 1% for an integrated luminosity of 1 fb^{-1}. Thus, the four-muon efficiency for this analysis is known with an absolute error of better than 4%.

The uncertainty on the muon p_T resolution directly propagates into the four-muon invariant mass $m_{4\mu}$ reconstruction while this almost does not affect the background distribution. The uncertainty on the muon p_T scale can be similarly calibrated from data using the measured J/ψ and Z peaks. The effect of these uncertainties on the number of background events in a signal window appears only on steep slopes of the $m_{4\mu}$ distribution. For the steepest part of the $m_{4\mu}$ distribution in the $180 - 200$ GeV/c^2 range, an uncertainty of $\delta b/b \approx 0.1\delta m_{4\mu}$, where $\delta m_{4\mu}$ is in GeV/c^2, is obtained. Therefore, in order to be able to neglect this effect, one needs to know the momentum scale with a precision of 0.1 GeV/c at $p_T \approx 50$ GeV/c. This can be easily achieved with just a few hundred $Z \to 2\mu$ events.

Uncertainties from misalignment

The limited knowledge about the exact position and orientation of the CMS tracker and muon subdetectors in the first data taking period affects the performances of the muon/track reconstruction and impacts on physics.

In order to study the impact of the mis-alignment of the CMS tracking devices on the tracking procedures, realistic estimates for the expected displacements of the tracking systems are supplied in several scenarios that are supposed to reproduce the mis-alignment conditions during the first data taking period corresponding to 10 pb^{-1} and 100 pb^{-1}.

The track reconstruction performance for what concern both the efficiency and the parameter resolution deteriorates with increasing misalignment. When the tracking system is aligned the p_T resolution of 100 GeV/c muons is about 1.5 – 2% in the barrel region, while it degrades to 2.5-4% and 5.5-8% in case of the 100 and 10 pb^{-1} scenarios respectively, as detailed in ref. [50]. The transverse impact parameter resolution is dominated by

the misalignment of the pixel detector degrading from 10 μm for the ideal geometry to values $\sim$ 15 and $\sim$ 50μm in case of the 100 and 10 pb^{-1} scenarios, respectively.

The width of the di-muon invariant mass distribution around the Z peak increases of about 12% for the 10 pb^{-1} misalignment scenario with respect to ideal detector conditions and is mostly related to the tracker misalignment compared to the muon chamber displacement effect.

Uncertainties from isolation

The main sources of efficiency losses and systematics on lepton isolation are the so-called underlying events (UE) originating from the same pp collisions, and the pile-up caused by random coincidences with other pp collisions.

The question of how well can the isolation cut efficiency be predicted on the basis of current Monte Carlo event generators, given the poorly known UE physics, was studied extensively in Ref. [51] in the case of 4μ final states. It was shown that the variation of the isolation cut efficiencies per muon with different UE models, can be as large as 5%. This result further depends on the details of the lepton isolation cut such as the minimum p_T threshold above which charged tracks are considered.

For this reason a strategy to measure the isolation cut efficiency using the experimental data themselves has been developed to have a good control of the systematic uncertainties with a data driven approach. This method, the so-called Random Cone Technique, was proposed and studied in details in the context of the CMS Physics TDR and in Ref. [51].

6.4 Event Skimming and Pre-Selection

The first step of the event selection is performed in two parts, called skimming and pre-selection. The main strategy of this step is to get rid of unuseful events, which will never pass the entire analysis selection, whilst preserving the maximal signal efficiency and the phase space for background systematic studies. The main aim of the first part, called skimming, is tech-

nical: to reduce event rate in order to preserve a manageable data volume. The main goal of the pre-selection is to eliminate fake events, in particular from QCD.

6.4.1 Event Skimming

A common skimming selection for all the three channels has been designed to select signal events with close to 100% efficiency, and reduce significantly backgrounds from QCD, W + jets and Z + jets (e.g. events containing jets recoiling against a Drell-Yan production of lepton pairs). In doing so this step also reduces in a lesser extend the other main backgrounds to this analysis: the $t\bar{t}$ and Z$b\bar{b}$. It is applied to all the electron and muon HLT paths (see Sec. 6.2) whose event rate, dominated by QCD, is around 30-40 Hz.

The skimming requires:

- at least two leptons (e or μ) with $p_T^l > 10$ GeV/c;
- one additional lepton with $p_T^l > 5$ GeV/c.

While the number of signal events is almost unchanged, the total event rate is reduced by a factor 50 with respect to the HLT selection (see first column of Fig.6.6). Figure 6.4 shows the skimming efficiency for the Higgs boson signal separately for the 4μ, $4e$, $2\mu 2e$ decay channels and for the combined $t\bar{t}$, Z$b\bar{b}$, ZZ backgrounds. The $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ channel is the decay mode with the higher efficiency: $\geq 99\%$ for $m_H \geq 135$ GeV/c^2. It was checked that none of the signal events that have been rejected by skimming would have passed the full event selection used for this analysis.

6.4.2 Event Pre-selection

After the Skimming, a set of pre-selection cuts is applied to suppress the contribution of "fake leptons". Dedicated selection requirements have been therefore adopted for each 4ℓ decay channel. In the following, the description of the analysis strategy will focus on the 4μ channel, where I could give

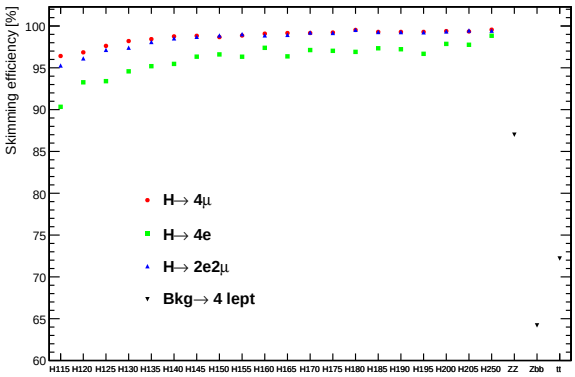


Figure 6.4: *Skimming efficiencies for all considered Higgs boson masses, for the three three 4-lepton channels and their main backgrounds.*

a greater contribution thanks to my knowledge of the reconstructed muon properties. The main aim of the pre-selection is to bring the QCD multijets and Z/W+jet(s) events to a level comparable to or below the contribution of the three main backgrounds, $t\bar{t}$, Z$b\bar{b}$ and ZZ, thus reducing the problem of the combinatorial ambiguities caused by the presence of more than 4 leptons.

The pre-selection of "signal-like" events contains four steps. The requirements for the 4μ channel are:

- at least two $\mu^+\mu^-$ pairs of identified muons with opposite charge and matching flavors. The muons are also required to satisfy $p_T^\mu > 5$ GeV/c in the barrel, $p_T^\mu > 3$ GeV/c and $P^\mu > 9$ GeV/c in the endcaps, and $|\eta^\mu| < 2.4$,
- at least two different matching pairs with invariant mass $m_{\mu^-\mu^+} > 12$ GeV/c^2,
- at least one combination of two matching pairs of muons with an invariant mass greater than 100 GeV/c^2,

- at least 4 loose isolated muon candidates.

The loose cut applied on the muon isolation observable (μISO, see Sec 6.5.1) is $\mu\text{ISO} < 60$. These requirements preserve the 99% of the signal events passing the previous pre-selection step, for $\text{m}_H = 150$ GeV/c^2. The requirement of two muon pairs with opposite charge and matching flavours is a restriction beyond skimming which brings additional rejection power against fake leptons. The cut on $m_{\mu^-\mu^+}$ protects against the contamination from low mass hadronic resonances while requiring that at least one Higgs candidate built from the contamination of 4 muons in a given event fulfills an invariant mass of 100 GeV/c^2 further suppresses unwanted events and brings a safe reduction of the phase space towards the signal phase space. Finally, imposing a loose isolation of muons is mandatory to eliminate the remaining fake QCD backgrounds.

The efficiencies of the different pre-selection steps for the 4μ channel are shown in Fig. 6.5 as a function of the Higgs boson mass. Note how the pre-selection efficiency reaches a plateau around $\text{m}_H = 170$ GeV/c^2 with a value $\sim 86\%$. The main loss of signal events comes from the requirement of at least four basic reconstructed muon objects within acceptance. The residual loss comes from a combination of the requirements of matching pairs of identified muons and of the loose muon isolation.

The suppression of the background rate obtained from the pre-selection steps is shown in Fig.6.6. The background events, largely dominating at the first steps by contributions involving multiple fake leptons (e.g. from light quark jets), are brought down to a level where $t\bar{t}$ and $\text{Z}b\bar{b}$ becomes important. In particular the fake contribution from QCD multijets is essentially eliminated after having imposed loose isolation.

The invariant mass reconstructed with four muons after pre-selection is shown in Fig. 6.7. In the case where more that four identified muons are found in a given event, the ambiguities have been resolved by building a single Higgs boson candidate from the pair of reconstructed vector bosons in the following way:

- the pair (i.e. with matching lepton flavour and opposite charges) with the invariant mass the closest to the Z mass,

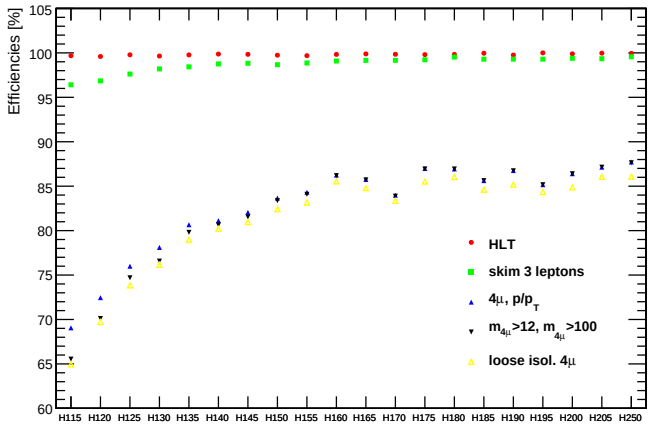


Figure 6.5: *Efficiency of HLT, skimming and each pre-selection step for $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ channel as a function of the Higgs mass.*

- and the next pair involving the highest p_T muons among remaining ones.

This ambiguity resolving criteria will be applied when building masses for the final event selection (Sec.6.6).

Finally, Table 6.3 shows the expected number of events per fb^{-1} in the 4μ channel after the pre-selection, for three Higgs boson masses and for the remaining backgrounds.

At this stage, the ZZ background does not emerge. The dominating background after pre-selection is the Z$b\bar{b}$ followed by the $t\bar{t}$ +jet(s) and the Z+jet(s) ones. All of these (and most notably the Z+jet(s)) will be strongly reduced in the final selection by tighter isolation, vertexing requirements and kinematic cuts.

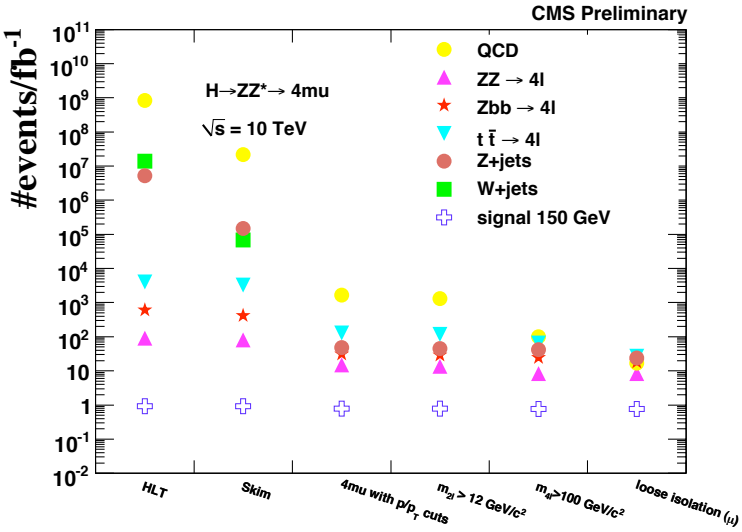


Figure 6.6: *Reduction of the event rate for QCD, Z/W + jet(s), $t\bar{t}$ +jet(s), $Zb\bar{b}$ and ZZ backgrounds, and $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$ signal at $m_H = 150$ GeV/c^2, after the skimming and each pre-selection step.*

6.5 Lepton Discriminating Observables

As showed in the previous section, the event sample obtained following pre-selection, in particular for the 4μ channel, is dominated by the sum of the $Zb\bar{b}$ and $t\bar{t}$ backgrounds, although the $t\bar{t}$ is suppressed well below the $Zb\bar{b}$. Henceforward, for the event selection, the focus is put on the reduction of leptons coming from the decays of the b quarks (in the $t\bar{t}$ sample they are produced from $t\bar{t} \rightarrow \mathrm{W}^{+}b\mathrm{W}^{-}\bar{b}$). Such leptons are likely to be accompanied by hadronic products from the fragmentation and decay processes initiated in the b-quark jets. Moreover, because of the long lifetime of b-hadrons, they are likely to have a large impact parameter with respect to the primary vertex. Thus, lepton isolation and lepton impact parameter measurements

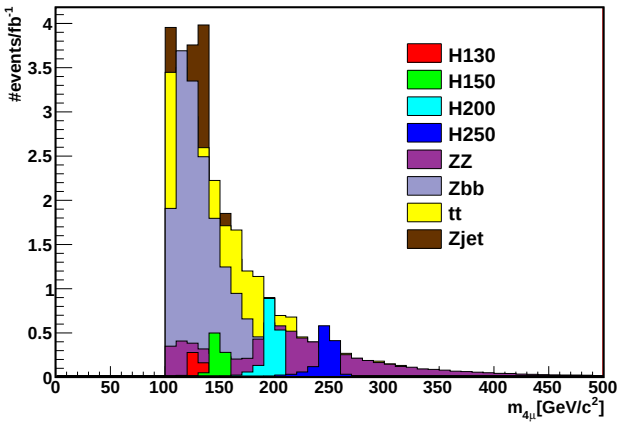


Figure 6.7: *Four leptons invariant mass after pre-selection in the* 4μ *channel for Higgs signal events and the backgrounds.*

should allow for a powerful rejection. While these characteristics might be sufficient to eliminate the leptons from heavily boosted b-quark jets in $t\bar{t}$ events, the b-quark jets in $Zb\bar{b}$ events are in general less collimated in the detector and lead to leptons with a softer p_T spectrum. In order to best preserve the signal detection efficiency while acting on low p_T^ℓ lepton candidates to suppress the $Zb\bar{b}$ background, the isolation criteria for the leptons from the pair at lowest $m_{\ell^-\ell^+}$ can be made p_T^ℓ dependent.

In the following (Sec. 6.5.1 to 6.5.3) the main discriminating muon observables for the event selection are described. These are:

- isolation,
- impact parameter,
- p_T of all four muons,
- two-muons invariant masses,
- four-muons invariant mass.

	number of events per fb^{-1}
$H \to ZZ^{(*)} \to 4\mu$ $m_H = 130$ GeV/c^2	0.47
$H \to ZZ^{(*)} \to 4\mu$ $m_H = 150$ GeV/c^2	0.86
$H \to ZZ^{(*)} \to 4\mu$ $m_H = 200$ GeV/c^2	1.70
ZZ	8.0
Z$b\bar{b}$	18.0
$t\bar{t}$ +jet(s)	25.8
Z +jet(s)	22.0
W +jet(s)	0

Table 6.3: *Summary of the number of events expected per fb^{-1} after the pre-selection in the 4μ channel, for three different Higgs boson masses and the backgrounds.*

An optimal combination of these quantities for a cut-based event selection will be discussed in section 6.6. Similar studies have been performed for electrons; further details on this subject can be found in the complete 4ℓ analysis paper [39].

6.5.1 Isolation

The four muons coming from the Higgs boson decay should appear isolated (i.e. not inside a jet). This provides an excellent way to distinguish the signal from the Z$b\bar{b}$ and $t\bar{t}$ backgrounds, where two muons are produced inside the b-jets.

The degree of isolation of a muon can be quantified by considering the energy or momentum of the particles in a cone around the muon track. Among the several observables investigated, the one that showed the highest discriminating power, in terms of the best background rejection for a high signal acceptance, is μIso, defined as

$$\mu\mathrm{Iso} = 2 \cdot \mu\mathrm{Iso}_{\mathrm{track}} + 1.5 \cdot \mu\mathrm{Iso}_{\mathrm{ECAL}} + \mu\mathrm{Iso}_{\mathrm{HCAL}} \ .$$

The tracker based and calorimeter based isolation variables have been de-

fined in the section 6.3.

Dedicated studies have been carried on the stability of the isolation observable; to check that the efficiencies do not strongly depend on the variable definitions.

After having measured the degree of isolation of all the leptons, a decision must be taken about the whole event. It has been found that the most effective way to define an event as 'isolated', is to consider the sum of the two least isolated muons out of the four ones. This sum $\mu\text{Iso}_{\text{2least}}$ is shown in the left part of Fig. 6.8 for the signal and main backgrounds.

The improvement in background rejection power is small (with respect to using only the least isolated muon) and mostly compatible within statistical fluctuations. But the choice make sense given that both b and $\bar{b}$ quarks tend to decay semileptonically, yielding two muons that are likely to be non-isolated. The signal efficiency for a cut on $\mu\text{Iso}_{\text{2least}}$ are shown in Figure 6.8 (right side) as a function of the background efficiency. For a cut $\mu\text{Iso}_{\text{2least}} < 30$ is 95%, while background efficiencies are 29% and 9% for $Zb\bar{b}$ and $t\bar{t}$ respectively. Efficiencies are calculated with respect to preselection.

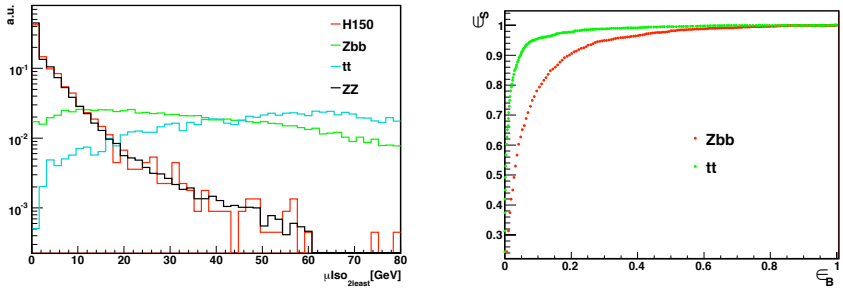


Figure 6.8: *(L) Muon combined isolation variable distribution for the sum of two least isolated muon, for signal and main backgrounds. (R) Combined isolation cuts power against $t\bar{t}$ and $Z\,b\bar{b}$ backgrounds.*

Another powerful selection criterion arises from observing the two-dimensional distribution of the isolation discriminating observable versus the p_T of the third or fourth lepton (sorted by decreasing order of p_T).

The $\mu\text{Iso}_{2\text{least}}$ versus the p_T of the third muon $p_{T,3}$ is shown in Fig 6.9, for the Higgs boson signal and for the Z$b\bar{b}$ background.

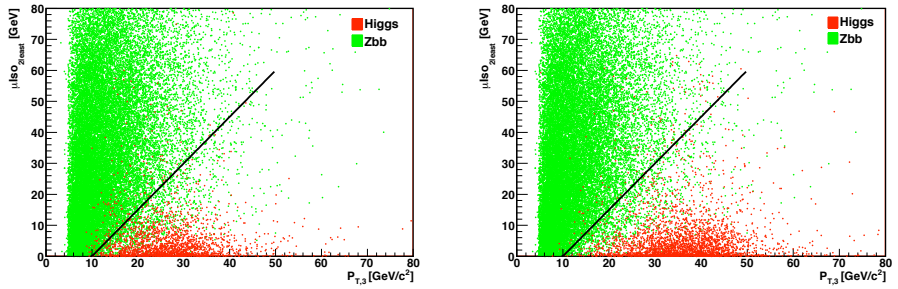


Figure 6.9: *Combined muon isolation variable versus p_T of the third lepton for m_H = 150 GeV/c^2 (L) and m_H = 200 GeV/c^2 (R).*

The signal and the background are seen to be well separated, so that the plane can be divided into two regions respectively dominated by the signal or the background. This conclusion has important consequences for the control of background systematics (see section 6.7). This separation can be expected, given that the leptons originating from the b-jets have usually low p_T, while those muons from Z (or W) decays are more energetic. Therefore in Z$b\bar{b}$ (and $t\bar{t}$) events, unlike signal events, the third and fourth muons are usually characterized by low p_T and worsen isolation. The signal and background regions are best separated by a slanting line of the form $\mu\text{Iso}_{2\text{least}} = A \cdot p_{T,3} - B$. The values of A= 1.5 GeV^{-1} and B = 15 are found to be optimal since they provide the best background rejection and signal maintenance for a Higgs boson mass of 150 GeV/c^2. Such a cut is used for all Higgs boson masses in this analysis, but could eventually be made m$_H$ dependant, as could be inferred from Fig 6.9.

6.5.2 Impact Parameter

The leptons from the Higgs boson signal originate from a common primary vertex in contrast to the leptons from at least one $\ell^+\ell^-$ pair reconstructed in $t\bar{t}$ and $Zb\bar{b}$ background events. The impact parameter (IP) for the signal leptons is expected to be compatible with zero. For the background, due to the long b-quark lifetime, the particles from b-jets usually have displaced production vertices and thus a large IP at the event primary vertex. This information is exploited to further improve the separation of the signal and background events.

To take into account the finite resolution of the detector, the variable $S_{IP} = \frac{IP^{3D}}{\sigma_{IP}^{3D}}$ is used for the analysis. The IP_{3D} is the distance between the primary vertex of the event and the point of closest approach of the track to this primary vertex, and σ_{IP}^{3D} the associated uncertainty. After sorting the S_{IP} of the four muons in increasing order, the fourth (i.e. worst) or the third (i.e. second worst), or both, can be used to distinguish signal from background.

The S_{IP} distribution for the fourth and the third muon in 4μ events is shown in the left part of Fig 6.10. The best criterion found for this channel is to require $S_{IP}(4^{th}\,\mu) < 12$ and $S_{IP}(3^{rd}\,\mu) < 4$. The power of the S_{IP} cuts used against $t\bar{t}$ and $Zb\bar{b}$ backgrounds after preselection, is illustrated in Fig. 6.10 (right side).

6.5.3 Kinematics

Taking advantage of the expectation of a narrow resonance in the $m_{4\mathrm{l}}$ spectrum, and of the likely presence of a real Z boson in the final state, the selection can be further improved using mass dependent kinematic requirements.

First, leptons coming from Z decays have a much harder p_T spectrum respect to those coming from b-jets. Figure 6.11 shows the transverse momentum spectrum of the four muons, sorted in decreasing order of p_T, for the $m_H = 150$ Gev/c^2 sample.

Second, the mass spectrum of the Z, Z$^{(*)}$ bosons distinguishes the Higgs

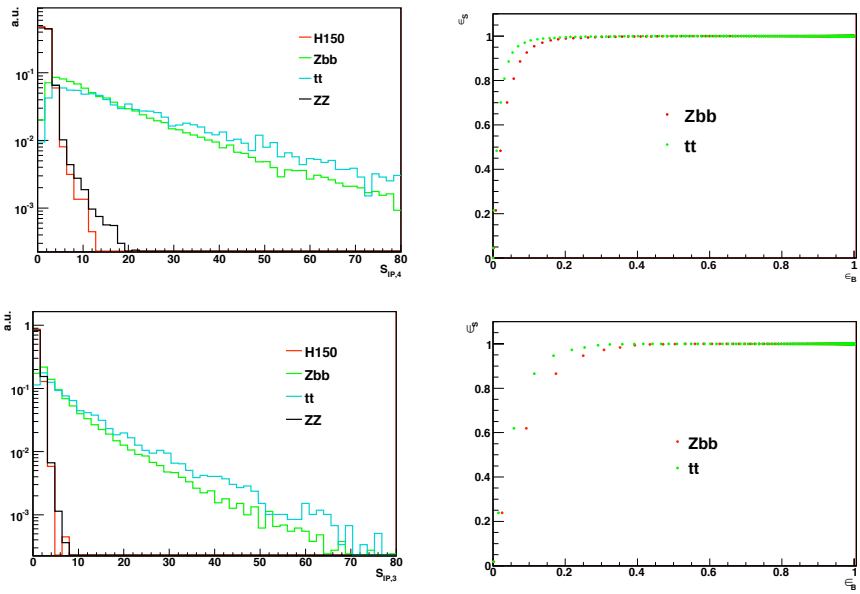


Figure 6.10: *L: Muon impact parameter significance S_{IP} distribution for the fourth (top) and the third (bottom) muon, for signal and main backgrounds. R: Power of the impact parameter significance cuts on the fourth (top) and the third (bottom) muon against $t\bar{t}$ and $Z\,b\bar{b}$ backgrounds.*

boson signal from its main background. The histograms of Fig. 6.12 highlight this spectrum separation.

The final event selection described in section 6.6 makes use of these requirements.

6.6 Results

The lepton transverse momentum, isolation and impact parameter observables, the two-lepton invariant masses and the four-lepton invariant mass $m_{4\ell}$

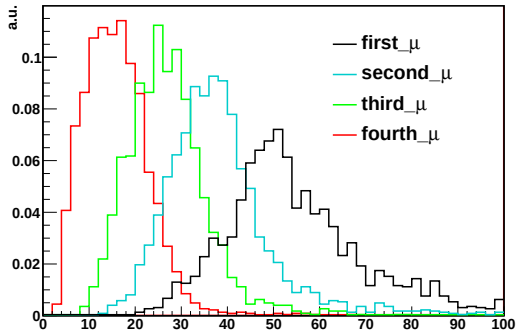


Figure 6.11: *Transverse momentum of the four muons sorted in decreasing order of p_T for the signal sample with M_H=150 GeV/C^2.*

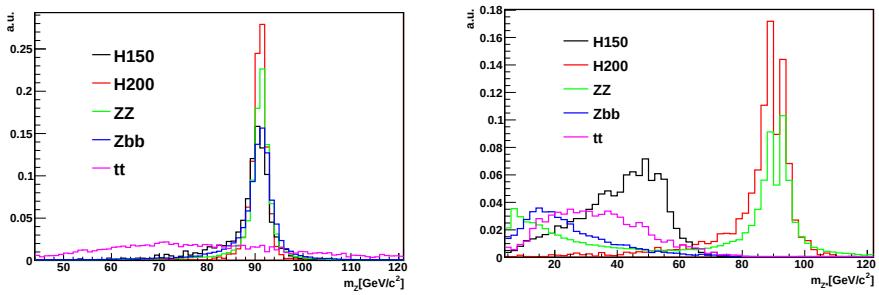


Figure 6.12: *Distribution of the Z (left) and $Z^{(*)}$ (right) invariant masses for different samples of the Higgs signal and its main backgrounds.*

can be combined to optimize the sensitivity to the Higgs boson as a function of the mass hypothesis m_{H}, and for a given integrated luminosity. Such mass dependent cut-based analyses been discussed in previous studies [52, 53] in the context of measurements at integrated luminosities of 30 fb^{-1}

at the LHC. For the start-up integrated luminosity of 1 fb^{-1} considered in this analysis, and given an improved suppression of the distinguishable background sources, it is found sufficient to consider a baseline cut-based selection, leaving only a sliding window cut in the measured $m_{4\ell}$ spectrum to optimize the sensitivity for a Higgs boson of given mass m_{H}. This allows for a simple search procedure covering the mass range from 115 GeV/C^2 to 250 GeV/c^2.

6.6.1 Event selection

A best set of discriminating observables, each offering high background rejection power, are combined in a baseline selection, designed to offer the best significance for the observation of a Higgs boson with a mass value around $m_{\text{H}} \simeq 150$ GeV/c^2.

The cuts are optmized separately for the searches in each of the 4ℓ final states. The baseline selection for the 4μ channel is summarized in Table 6.4. The variable μIso represent the sum of the two least isolated muons, as explained in Sec. 6.5.1.

Isolation	$\mu\text{Iso} < 30$ $\mu\text{Iso} < 1.5 \cdot p_T^3 - 15$ $\mu\text{Iso} < 2.0 \cdot p_T^4 - 10$
IP	$s^{\mu}_{IP}(4^{th}) < 12$ & $s^{\mu}_{IP}(3^{rd}) < 4$
Lepton p_T	$p^{\mu}_{T\,min} > 5$ GeV/c
M_Z	[50 GeV/c^2, 100 GeV/c^2]
$M_{Z^\star}$	[20 GeV/c^2, 100 GeV/c^2]

Table 6.4: *Set of baseline selection cuts for the 4 muons final state.*

The four muon reconstructed invariant mass spectrum after baseline selection is shown in Fig. 6.13. The $t\bar{t}$+jets background is completely eliminated. The $\text{Z}b\bar{b}$ background is considerably reduced and now only survives towards low masses, with an event rate far below that of the ZZ continuum.

The SM Higgs boson signal expected on average for an experiment at

1 fb^{-1} of integrated luminosity is superimposed for illustration in Fig. 6.13 for various mass hypothesis. The signal is observed as a narrow peak. The mean number of events expected from the signal is seen to be comparable or larger than that expected from the background in a narrow mass window centered on the signal. Thus, after the m_{H} independent baseline selection, a signal emerges above the background for any particular m_{H} value, over the full mass range.

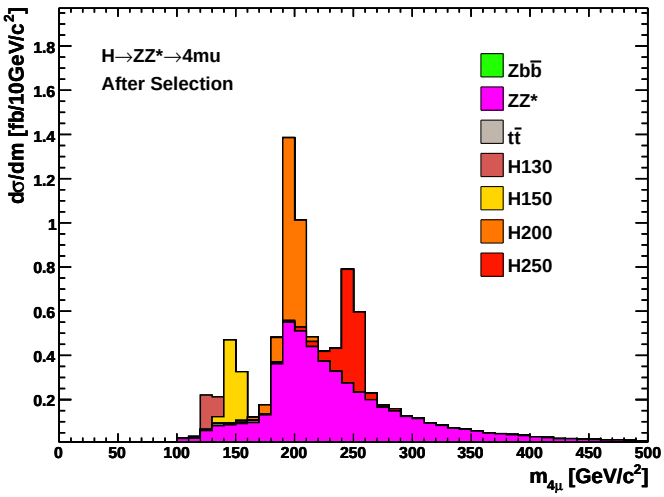


Figure 6.13: *Four muon reconstructed invariant mass spectrum after baseline selection. The spectrum should be interpreted as the mean spectrum expected for an experiment at 1 fb^{-1} of integrated luminosity. The 4μ spectrum is composed mostly of the $\text{ZZ}^{(*)}$ continuum and it receives an instrumental contribution from $\text{Zb}\bar{\text{b}}$. Superimposed here for illustration is the expectation from a Higgs boson signal for mass hypotheses of 130, 150, 200 and 250 GeV/c^2.*

6.6.2 Analysis sensitivity to the Higgs signal

In order to quantify the sensitivity of the experiment to the presence of a Higgs boson signal, a simple counting experiment approach is used. For each possible m_{H} hypothesis, the events are counted in a reconstructed mass window $m_{4\ell} \pm 2\sigma_{m4\ell}$. The reconstructed mass $m_{4\ell}$, and hence the central value for the mass window, slightly differs (by less than about 0.5%) from the true Higgs boson mass m_{H}, given small and unavoidable non-Gaussian fluctuations of the reconstructed mass towards low masses. For simplicity the width $\sigma_{m4\ell}$ is nevertheless taken here from a Gaussian fit to the signal distribution for each given m_{H} hypothesis. The mass window slides along the measured mass spectrum of Fig. 6.13 to test the various possible m_{H} hypotheses.

The mean expected number of observed (signal plus background, N_{s+b}) and expected background only (N_b) events evaluated at $L = 1$ fb^{-1} in the $m_{4\ell} \pm 2\sigma_{m4\ell}$ mass window for a selected set of m_{H} masses are given in Table 6.5.

For a counting experiment, the log-likelihood ratio $-2lnQ$ is given by the following equation:

$$-2lnQ = -2\,(N_{s+b})\,\ln\,(1 + N_s/N_b) + 2\,N_s. \tag{6.1}$$

The blue points in Fig. 6.14 show the mean expected values for $-2lnQ$ in assumption of a background-only hypothesis. If one sees in real data large negative values, it could be interpreted as a signal of a Higgs boson. Red points indicate where one may expect to find data points, should the Higgs boson exist at one of these masses. Fig. 6.14 does not include systematic errors.

The actual quantitative measure for an event excess is a significance S_{cP} [54]. To evaluate it, a probability for the background to fluctuate to number of events equal or greater than the observed number N_{s+b} is first calculated. The calculation of such a probability includes Poisson fluctuations and uncertainties on the level of background. Then, this probability is converted in an equivalent number of one-sided tail standard errors ('sigmas') of the Gaussian distribution. The mean expected sensitivity for the

m_H (GeV/c^2)	$Reco(m_{H_m} - m_{H_M})$ (GeV/c^2)	N_{s+b}	N_b
115	112.34 - 116.88	0.036	0.009
120	117.43 - 122.14	0.069	0.014
125	122.30 - 127.01	0.125	0.023
130	127.24 - 132.18	0.246	0.033
135	132.15 - 137.29	0.353	0.043
140	137.02 - 142.31	0.457	0.044
145	141.74 - 147.36	0.544	0.052
150	147.02 - 152.32	0.547	0.048
155	151.59 - 157.57	0.521	0.051
160	156.36 - 162.65	0.319	0.065
165	161.27 - 167.64	0.182	0.062
170	165.85 - 173.12	0.200	0.076
175	170.71 - 178.24	0.258	0.089
180	175.46 - 183.26	0.445	0.130
185	180.40 - 188.52	1.032	0.265
190	184.88 - 193.87	1.540	0.430
195	188.79 - 199.22	1.696	0.543
200	193.88 - 204.60	1.722	0.550
205	198.74 - 209.57	1.623	0.529
250	239.32 - 258.02	1.368	0.491

Table 6.5: *Mean expected number of observed (signal plus background) and expected background only events at $L = 1$ fb^{-1} for the 4μ final state. Events are counted in a $m_{\mathrm{H}} \pm 2\sigma_m$ windows around Higgs boson reconstructed masses, which slightly differs (by less than 0.5%) from the true Higgs boson mass m_H.*

4μ channel is given numerically in Table 6.6 and shown in Fig. 6.15: it clearly does not exceed 1.4σ in the whole m_H spectrum, mainly because of the low statistic available with 1 fb^{-1}.

In the absence of a significant deviation from the background-only hypothesis, a 95% confidence limit on the $\mathrm{H} \rightarrow \mathrm{ZZ} \rightarrow 4\mu$ cross-section σ in

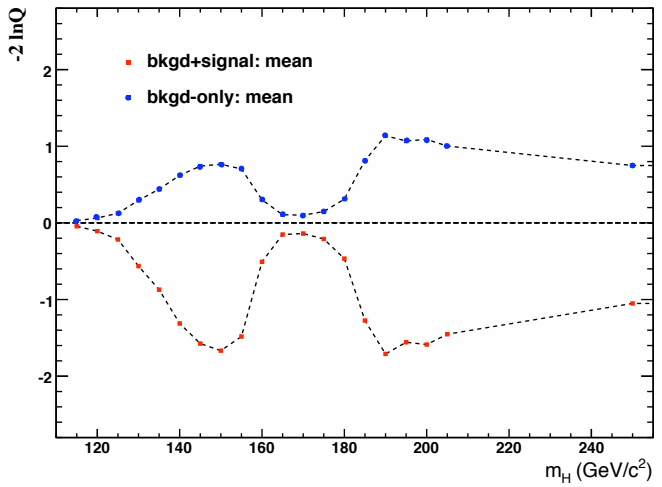


Figure 6.14: *The expected $-2lnQ$ values for background-only and signal+background hypotheses in the 4μ final state.*

terms of the ratio $r = \sigma/\sigma_{\mathrm{SM}}$ with respect to the SM Higgs boson cross section σ_{SM} is set up. Using the Bayesian approach [55], the exclusion limit is derived from the posteriori likelihood function for the ratio r, given that n_0 events are observed, while the expected background is b and the Standard Model Higgs signal is s_{SM}:

$$L(r) = \frac{1}{A} \int p\,(\,n_0\,|\,x\,b + y\,(r\,s_{\mathrm{SM}})\,)\;\rho_b(x)\,\rho_s(y)\,dx\,dy, \qquad (6.2)$$

where $p(n_0|\mu)$ is the Poisson probability function, x and y are modifying factors (systematic errors) for the expected background b and signal s event counts; their probability density functions being $\rho_b(x)$ and $\rho_s(y)$. A stands for the normalization factor.

The 95% C.L. limit on the ratio r is set by finding a solution for the following equation:

$$\int_0^r L(r)\,dr = 0.95 \tag{6.3}$$

The experimental uncertainties on the signal (with the same amount as for the ZZ background) and background, including luminosity, are taken to be 100% correlated, which, being a very good approximation in this analysis, is also the most conservative assumption in terms of the resulting limits. The modifying factor *pdf* $\rho(x)$ is taken to be a log-normal distribution. At the end, the effect of including systematic errors is rather small.

The mean expected ratio $< r >$ for the background-only hypothesis is calculated as follows:

$$< r >= \sum_{n_0=0}^{\infty} r(n_0) \cdot p(n_0|b) \tag{6.4}$$

The results are given in Table 6.6 for various m_{H} hypothesis and expressed in terms of the ratio of excluded and SM cross-sections $R_{95\%C.L.} = \sigma_{95\%C.L.}/\sigma_{SM}$. These results are represented also in Fig. 6.16. One can see that the statistics available with 1 fb^{-1} of data is not enough to exclude the SM-like Higgs at 95%C.L. in the Higgs mass range considered.

6.6.3 Combined Results

The same variables described in the previous section for the 4μ channel are here presented combining the results of all the three 4ℓ final states. This has been done in order to exploit the full potential of the leptonic decays of the $H \rightarrow ZZ^{(*)}$ process, which have in common a quite clear experimental signature but a very low event rate if compared with the background amount at the LHC. Figure 6.17 shows, in order from the top to the bottom, the log-likelihood ratio $-2lnQ$, the significance and the mean expected ratio $< r >$ for the background only hypothesis after the matching of all the 4ℓ analysis with 1 fb^{-1} of data. The background rejection power clearly increases, even if the the total significance does not exceed 2.5 in the whole Higgs mass range. Table 6.7 summarizes the same results in a numeric view.

m_H (GeV/c^2)	Significance	95 % C.L. for R
115	0.211	88.2
120	0.329	46.4
125	0.465	25.7
130	0.751	13.1
135	0.933	9.3
140	1.146	7.2
145	1.255	6.0
150	1.291	6.0
155	1.218	6.2
160	0.713	10.4
165	0.389	18.5
170	0.373	16.8
175	0.458	12.8
180	0.683	7.6
185	1.129	3.5
190	1.308	2.5
195	1.248	2.4
200	1.260	2.3
205	1.205	2.4
250	1.025	2.9

Table 6.6: *Significance and values of R at 95% confidence level for selected Higgs boson masses expected for the 4μ final state.*

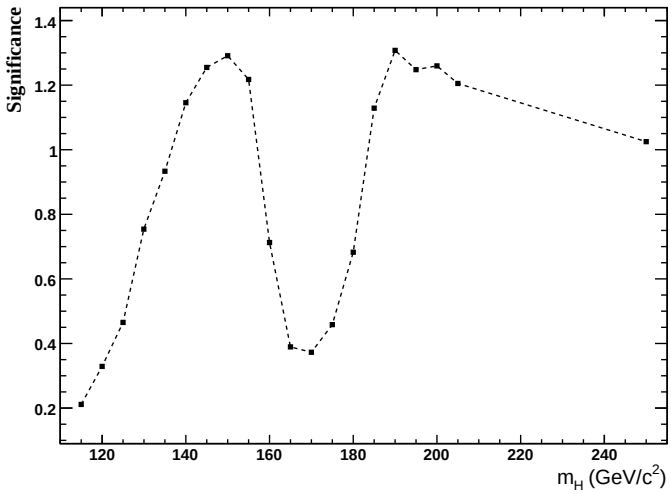


Figure 6.15: *The expected significance of an event excess in assumption of a SM Higgs boson presence for the 4μ final state.*

6.7 Control of Backgrounds from Data and Systematic Uncertainties

For a very low integrated luminosity at the LHC, while the discovery of a SM-like Higgs bosons in the $H \to ZZ^{(*)}$ channel remains unlikely, a particular emphasis will be put on the understanding of detector measurement uncertainties as well as on the control of background rates and uncertainties from data.

After the baseline sequential set of cuts (Sec.6.6) which includes a tightening of the isolation requirements with respect to the pre-selection, the $Zb\bar{b} \to 4\ell + X$ and $t\bar{t} \to 4\ell + X$ backgrounds are very much reduced over the full $m_{4\ell}$ range of interest. The main remaining background is the undistinguishable $ZZ^{(*)} \to 4\ell$ continuum. Hence, it is of outmost importance to obtain an estimate for this contribution from data. At 1 fb^{-1} of integrated luminosity, only a handfull of events is expected in the signal region. Thus,

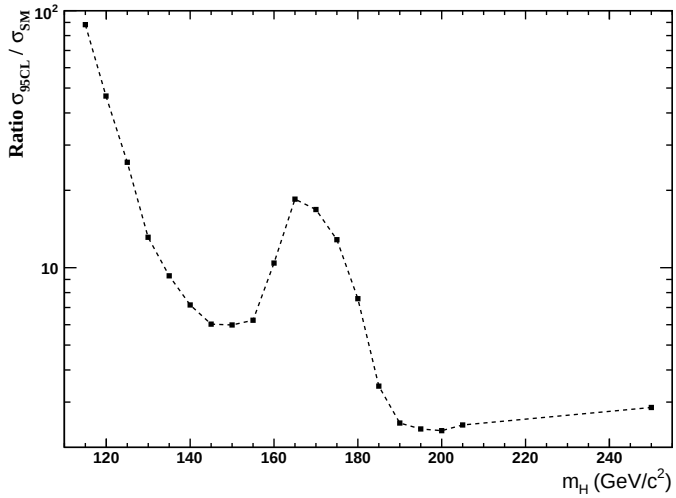


Figure 6.16: *$H \to ZZ \to 4\mu$: Exclusion limits on the ratio r for a range of Higgs mass points. The ratio $r_{95CL} = \sigma_{95CL}/\sigma_{SM}$ is the ratio between the cross section σ_{95CL} excluded with 95% confidence level and that of the Standard Model Higgs boson production, σ_{SM}. For the mass regions where $r_{95CL} < 1$, the SM Higgs boson would be excluded, if indeed it is not there.*

overall, the systematic uncertainties affecting the significance of the observations are expected to be largely dominated by statistical errors.

In this section the measurement and control of main backgrounds is presented, following an usual procedure which consist of choosing the control region outside the signal phase space and then verify that the event rate changes according to our expectations. A great care should be taken when choosing the control region, since two reducible backgrounds $Zb\bar{b} \to 4\ell + X$ and $t\bar{t} \to 4\ell + X$, after relaxing some cuts, quickly become dominant making this extrapolation more difficult. The main sources of experimental systematics are summarized at the end this section (6.7.3) and the summary of uncertainties for the measurement of the $ZZ^{(*)}$ background is given.

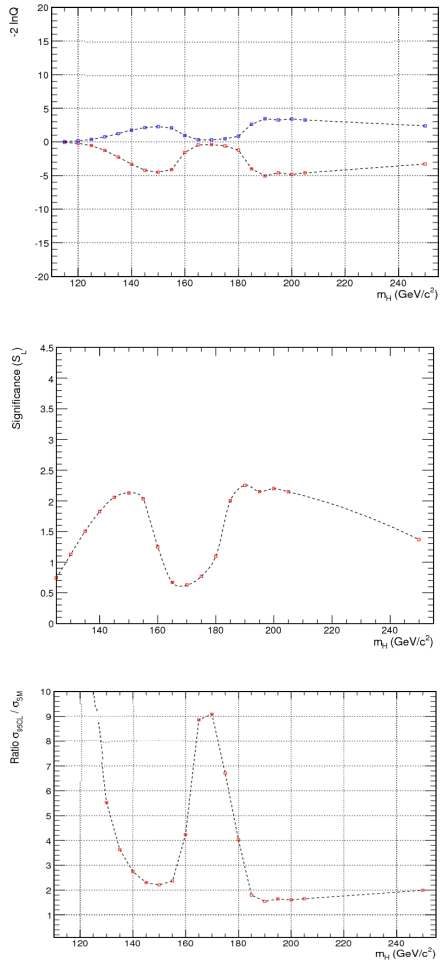


Figure 6.17: *Combined 4ℓ results. Top: The expected $-2lnQ$ values for for background-only and signal+background hypotheses. Center: The expected significance of an event excess in assumption of a SM Higgs boson presence. Bottom: Exclusion limits on the ratio between the cross section σ_{95CL} excluded with 95% confidence level and that of the Standard Model Higgs boson production, σ_{SM}.*

m_H (GeV/c^2)	N_{s+b}	N_b	Significance	95 % C.L. for R
115	0.08	0.04	0.30	42.66
120	0.16	0.08	0.46	20.70
125	0.33	0.12	0.74	10.06
130	0.61	0.15	1.13	5.53
135	0.95	0.18	1.50	3.64
140	1.26	0.20	1.83	2.74
145	1.52	0.21	2.06	2.30
150	1.58	0.21	2.13	2.21
155	1.47	0.20	2.03	2.37
160	0.83	0.24	1.25	4.24
165	0.40	0.25	0.67	8.86
170	0.40	0.29	0.63	9.09
175	0.55	0.36	0.77	6.71
180	0.99	0.54	1.10	4.03
185	2.59	0.99	2.00	1.80
190	3.61	1.60	2.25	1.54
195	3.64	1.87	2.15	1.64
200	3.82	1.97	2.20	1.60
205	3.72	1.98	2.15	1.65
250	2.83	1.65	1.81	1.99

Table 6.7: *Mean expected number of observed (signal plus background) and expected background only events from the combined $H \rightarrow ZZ \rightarrow 4\ell$ channels at $L = 1\ fb^{-1}$. Events are counted in a $m_{\rm H} \pm 2\sigma_m$ windows around Higgs boson reconstructed masses.*

6.7.1 Measurement of the ZZ background from data

Direct measurement.

After all analysis cuts are applied, the $ZZ \rightarrow 4l$ process comprises the major "irreducible" background. Therefore the knowledge of the number of attended events from this background and its 4ℓ invariant mass shape

became crucial to claim any possible Higgs evidence.

The strategy to measure the ZZ background directly from data starts from the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ analysis cuts and lepton discriminating variables. Among the leptons from Zbb and tt processes there are in fact at least one or two of them with low transverse momentum and worsen isolation respect to the leptons coming from the ZZ background, which has a kinematic configuration of the final state quite similar to the $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ process. This because leptons originating from b-decay are inside jets and have low p_T while those from Z (or W) decay are at the same time more energetic and isolated. Moreover, due to the long life-time of the b quark, leptons from Zbb and tt decays have impact parameters larger than those from the ZZ background.

Accordingly with the above statements, the analysis for the ZZ search can use the same $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ event selection described in the previous paragraphs. The only different kinematic cut concerns the Z, $Z^{(*)}$ invariant masses, which are required to be in the ranges 80-100 GeV/c^2 and 70-110 GeV/c^2 respectively. This more stringent selection has been chosen in order select both the Z on shell.

The first plot of figure 6.18 shows the analysis significance (obtained combining all the 4ℓ final states) with respect to the luminosity needed after the selection cuts on the muon isolation and the impact point. In the the second and the third plots of Fig.6.18 the same correlation between the analysis significance and the luminosity needed is shown after the cuts on the isolation parameter versus the p_T of the third and the fourth lepton, respectively. Only 140 pb^{-1} of data will be needed for a re-discovery of the ZZ process.

My work focused on the search for simpler discriminating variables and kinematics cuts of the $ZZ \rightarrow 4\ell$ analysis without loosing in performance. These simpler cuts may be used with the few statistics available during the first LHC running period, holding the more refined analysis by a greater integrated luminosity. Moreover this simplified analysis will be flexible enough to run in the same way for all the 4ℓ channels. In the following the new discriminating parameters and the kinematics cuts are described for the 4μ final state. Results of the analysis significance obtained combining

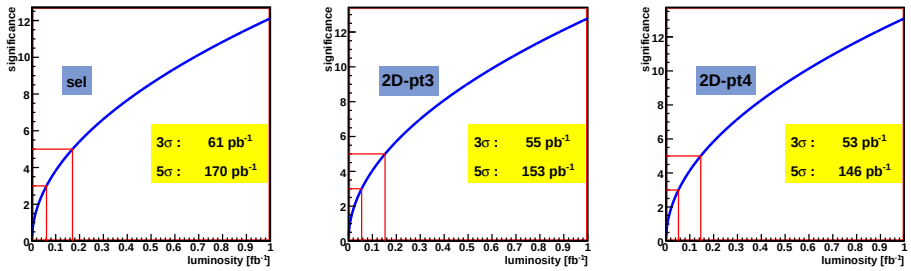


Figure 6.18: *Significance as a function of the luminosity needed after the one-dimensional selection cuts (L), the successive first two-dimensional cut (C) and the final two-dimensional cut (R).*

all the possible decay channels are shown too.

In order to simplify the definition of Sec.6.5.1, the isolation variable has been defined in this study as the track isolation only, $\mu\text{Iso}_{\text{track}}$. In Table 6.19 the one-dimentional cuts of the old selection and the attended events for each sample are compared with the new simpler version. Figure 6.20 shows the isolation parameter of the worst isolated muon and the distribution of the largest muon impact point. The black vertical lines correspond to the kinematic cut chosen for the event selection.

variable	**H→ZZ→4μ cuts**				**new cuts**			
muon isolation	μIso[0]+μIso[1]<30				$\mu\text{Iso}_{\text{track}}$[0]<13			
muon IP	IP[0]<12 , IP[1]<4				IP[0]<5			
N_{evt} for L=1fb^{-1}: samples:	4.735 ZZ	0.0276 Zbb	0 tt	0 Zjet	4.487 ZZ	0.030 Zbb	0.010 tt	0.001 Zjet

Figure 6.19: *Comparison on the one-dimentional cuts definition between the* $\text{H} \rightarrow \text{ZZ}^{(*)} \rightarrow 4\mu$ *analysis and the simplified one. The number of expected events for both selections is written in the bottom part of the table.*

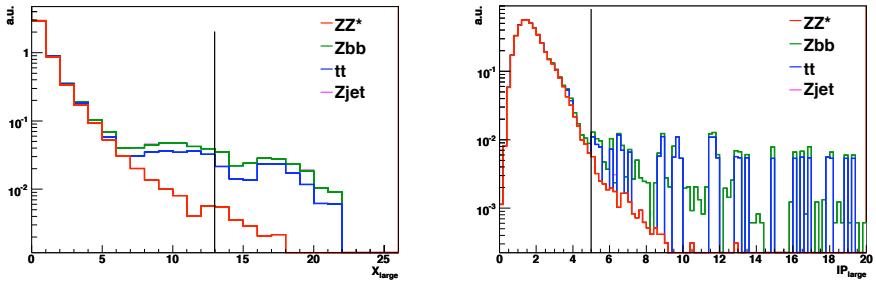


Figure 6.20: *Isolation (L) and Impact Point (R) distributions of the worst muons. Different samples have been superimposed. The black vertical lines correspond to the kinematic cuts chosen for the event selection.*

Table 6.21 summarizes the re-definition of the two-dimensional cuts and their impact on the final analysis selection. They have a very high backgrounds rejection power, leading the final significance to be higher than the old one. Figure 6.22 shows the evident kinematic separation between the ZZ and $Zb\bar{b}$ events.

Combining the present result with the other $ZZ \rightarrow 4\ell$ simplified analysis, the significance improves up to 13.4 with 1 fb^{-1} of data (see Fig.6.23), a value which is absolutely compatible with the one found with the more refined ZZ analysis.

Indirect measurement.

There are two data-driven methods that can be used to evaluate the level of the ZZ-background in a priori defined signal mass window for any given Higgs boson mass m_{H}: $m_{\mathrm{H}} - 2\sigma_m < m_{4l} < m_{\mathrm{H}} + 2\sigma_m$. One method calls for using sidebands in the m_{4l} spectrum outside of a signal region, while the other relies on measuring $Z \rightarrow 2l$ event count, which is closely related to the ZZ production ($q\bar{q}$-annihilation is the leading order diagram for both; next-to-leading order diagrams are very similar as well). In both

	H→ZZ→4μ cuts				**new cuts**			
	μIso[0]+μIso[1]<1.5*pt3-15				μIso_{track}[0]<0.37*pt3-3.7			
N_{evt} for L=1fb^{-1}: samples:	4.698 ZZ	0.0061 Zbb	0 tt	0 Zjet	4.6978 ZZ	0.0037 Zbb	0 tt	0 Zjet
ZZ significance:	**7.30**				**7.73**			
	μIso[0]+μIso[1]<2*pt4-10				μIso_{track}[0]<1.5*pt4-7.5			
N_{evt} for L=1fb^{-1}: samples:	4.614 ZZ	0.0031 Zbb	0 tt	0 Zjet	4.6710 ZZ	0.0012 Zbb	0 tt	0 Zjet
ZZ significance:	**7.65**				**8.34**			

Figure 6.21: *Comparison of the* $H \rightarrow ZZ^{(*)} \rightarrow 4\ell$ *two-dimensional analysis cuts and the simplified ones. The number of expected events for each sample and the significance after both the two selection cuts are shown too.*

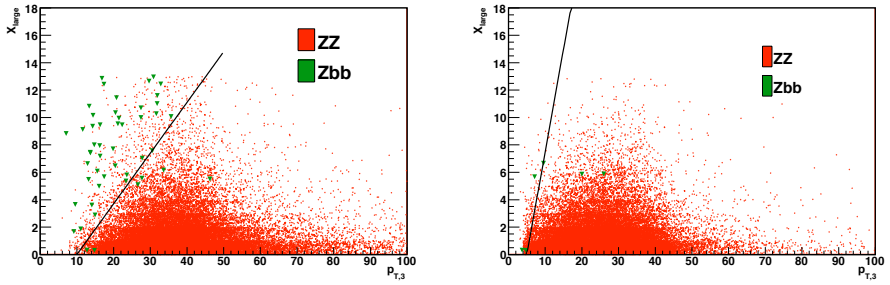


Figure 6.22: *Isolation as a function of the transverse momentum of the third (L) and the fourth (R) muon. The black lines correspond to the kinematic cut chosen for the event selection.*

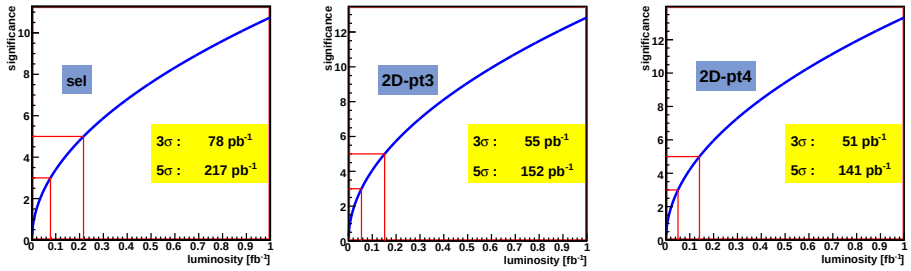


Figure 6.23: *Significance as a function of the luminosity needed after the one-dimensional selection cuts (L), the successive first two-dimensional cut (C) and the final two-dimensional cut (R) of the simplified analysis.*

cases, the background rate $N_{ZZ}(\Delta m)$ in any given signal mass window (symbolically Δm) can be predicted from the observed event count in the control region N_{CR}, i.e.

$$N_{ZZ}^{\text{predicted}}(\Delta m) = \rho(m_{\text{H}}) \cdot N_{CR}^{\text{measured}}. \tag{6.5}$$

The factor $\rho(m_{\text{H}})$ has the following structure:

$$\rho(m_{\text{H}}) = \frac{N_{ZZ}^{\text{theory}}(\Delta m) \cdot \epsilon_{ZZ}}{N_{CR}^{\text{theory}} \cdot \epsilon_{CR}}, \tag{6.6}$$

where ϵ's are the reconstruction/cut efficiencies.

Both methodologies have been studied in detail and the results are documented elsewhere [52, 56]. A brief summary of the results is given below.

The uncertainties for such predictions have two distinct components: systematic errors associated with the factor $\rho(m_{\text{H}})$ and statistical errors associated with measuring N_{CR}.

In the mass range 130-250 GeV/c^2, systematic errors on the factor $\rho(m_{\text{H}})$ arising from theoretical uncertainties (pdf and QCD scale) are 2-4% for the case of normalization to $Z \rightarrow 2l$ production and $< 3\%$ for

normalization to the $ZZ \to 4l$ sidebands. Both have been evaluated at the NLO level.

Systematic errors on electron/muon reconstruction and isolation cut efficiencies cancel out either partially (normalization to $Z \to 2l$, which has two leptons) or nearly completely (sidebands). The tag-and-probe methods will allow to measure the lepton reconstruction and isolation cut efficiencies with a 1%-uncertainties.

For both methods, the luminosity uncertainties cancel out completely by definition.

At L=1 fb^{-1} the sidebands method is strongly affected by the low statistics associated to the $ZZ \to 4l$ process as respect to the control sample $Z \to 2\ell$. It remains nevertheless a very important crosscheck since, unlike the method of normalization to Z, the use of sidebands guaranties a nearly perfect cancelation of uncertainties associated with the event reconstruction efficiencies.

Both methods presented rely on the knowledge of the overall ZZ background 4ℓ invariant mass shape, which can be measured with the method describe above, in the present sub-section.

6.7.2 Measurement of the $Zb\bar{b}$ background from data

A contamination of $Zb\bar{b}$ events remains after pre-selection, and a small amount of such events will survive the "baseline" m_{H} independent selection cuts of section 6.6 used for the final analysis, especially in the low m_{H} range. An appropriate strategy is therefore needed to control the event rate of such process from data.

The control of the $Zb\bar{b}$ event rate in the signal phase space profits from the strategy used at selection level to reject such process. Both tasks are fulfilled by finding a kinematic region with a high discriminating power between signal and $Zb\bar{b}$ background.

Two very discriminating variables are the isolation parameter, defined in section 6.5.1, and the transverse momentum of the third and fourth muons (sorted in decreasing order). As explained in Sec.6.6 and Sec. 6.7.1, the most effective way to suppress events form Zbb interactions while keep-

ing the Higgs and the ZZ efficiency almost untouched is to combined these variables. Therefore the rejected region (i.e. $\mu\text{Iso}_{2\text{least}} > 1.5 \cdot p_{T,3} - 15$) has been considered to study the $Zb\bar{b}$ background. Unfortunately this control region is populated not only by events from the $Zb\bar{b}$ process but also from $t\bar{t}$ and Z+jets interactions. My work focused on the search for an appropriate method to separate the $Zb\bar{b}$ sample from the remaining ones. The invariant mass of the first Z candidate has been used as discriminating variable since it can be fitted with the convolution of two gaussian functions obtaining a good separation between the $Zb\bar{b}$, Z+jet processes and the $t\bar{t}$ one. This because the $Zb\bar{b}$ and Z+jets show a real Z resonance decaying into two muons, while in $t\bar{t}$ events instead the two muons do not resonate. Further studies are on-going to find a discriminating variable also for $Zb\bar{b}$ and Z+jets, thus in this analysis they are assumed to be separable with other methods and the sytematics coming form the non-subtraction of the Z+jet sample are not considered.

To better study the estimate uncertainty on the background from data, a "pseudo-experiment" with 1 fb^{-1} of luminosity is performed.

The event contributing to the mass of the first Z candidate are illustrated in the left part of Fig.6.24, after having required $\mu\text{Iso}_{2\text{least}} > 1.5 \cdot p_{T,3} - 15$, while in the right part the same Z distribution with the superimposed fit (black line) and its components (red and green lines) are shown.

By integrating the two functions of the fit, the number of events from $Zb\bar{b}$ and Z+jet results 34 while from $t\bar{t}$ 43. The uncertainties are computed comparing the results of the fit with the attended events from Monte Carlo, resulting on a systematic error of 26% for $t\bar{t}$ and 15% for $Zb\bar{b}$ and Z+jets combined together.

Further pseudo-experiments with 1 fb^{-1} of data have been performed to check the dependence of the results from the analysis method. Poissonian variations of the initial number of events have been also considered. The systematic errors resulted to be around 30% for both $Zb\bar{b}$ plus Z+jets and $t\bar{t}$.

The uncertainties on the measurement of the isolation variable will be determined by the *tag and probe* method using high statistic $Z \to \ell\ell$ events

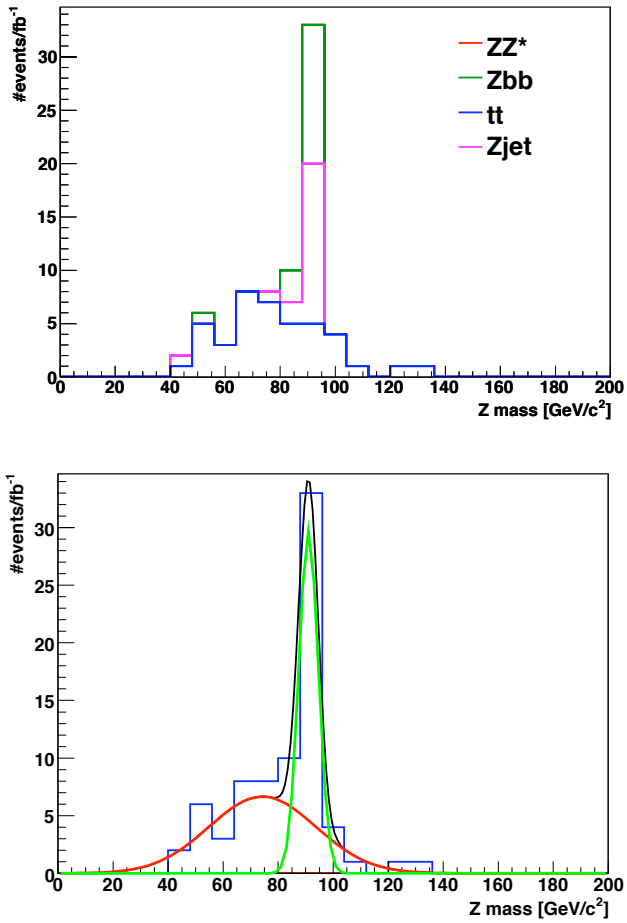


Figure 6.24: *Top: Attended event contributions from each sample of interest to the Z invariant mass spectrum in the region* $\mu\text{Iso}_{2\text{least}} > 1.5 \cdot p_{T,3} - 15$ *with 1* fb^{-1} *of data. Bottom: Total Z invariant mass spectrum with the superimposed fit (black line) and its components (red and green lines).*

and $pp \rightarrow b\bar{b}$ events (e.q. see [57]).

6.7.3 Experimental uncertainties

The main experimental sources of systematic uncertainties for the four lepton channels are:

- the integrated luminosity,
- the trigger efficiency,
- the lepton reconstruction and identification efficiency,
- the lepton isolation cut efficiency,
- the mis-calibration and mis-alignment,
- the four-lepton mass M_{4l} absolute scale and resolution $\sigma(m_{4l})$.

The absolute luminosity normalization can be obtained from the measurement of inclusive Z or W vector boson production with the production cross-section taken from the theory. A limitation on the precision of the integrated pp luminosity comes from the limited knowledge on parton density functions in the calculation of the theoretical cross-section. Otherwise, uncertainties on the detector precision at early stages of operations at the LHC will affect the cross-section obtained e.g. from a fit to the measured Z peak. The uncertainty on the integrated luminosity is expected to be of $\mathcal{O}(5\%)$ for measured luminosities of up to 1 fb^{-1}.

The uncertainty on the trigger efficiency is expected to be negligible in the signal-like phase space of the four-lepton analysis where the absolute trigger efficiency approaches 100 %.

Methods have been developed to evaluate experimental systematics from data where possible, and to make use of extrapolations from Monte Carlo simulation of the detector response where needed. Data-driven techniques are used for example to evaluate systematics on lepton isolation (see Sec.6.3.2). This is complemented by Monte Carlo extrapolation towards the very low p_T range. Monte Carlo simulations are used to estimate the

effect of of mis-alignment and mis-calibration on individual leptons and propagate it to the measurement of m_{2l} and m_{4l} taking into account the residual errors expected from low luminosity data analysis. Thus, comparing for example the Monte Carlo expectation for m_{2l} with the measured Z mass resolution, will help in establishing the credibility of the experimental systematic errors which are noticeably difficult to monitor at the early stage of the detector operation at the LHC, when the changes in the system conditions are frequent.

The lepton momentum scale has been studied in details in ref.[57] and the systematics effects on the Z cross sections have been computed. In this analysis for the Z decaying into muons, only "global muons" are considered, with very similar cut of the Z analysis and thus the systematics errors are negligible (i.e. 0.05%).

Summary

This thesis describes the various steps of the muon detection and reconstruction in the context of the CMS experiment starting from the single hit measurement in the Drift Tubes to the characterisation of the complete muon reconstruction. This "reconstructed muon" contains all the information needed within various physics analyses, as the Higgs search.

Drift Tubes are detectors based on time measurements. Therefore they have to be properly calibrated in order to translate temporal information into hit positions in the detector. Moreover the experimental noise inside a given cell may affect the calibration procedure if no correctly understood. Then the so-called "Local Reconstruction" combines all the hits together in a chamber and builds a muon track segment. Drift Tubes are designed to provide these track segments to the muon reconstruction which starts at the level of each individual chamber inside the CMS spectrometer.

I studied the behaviour of the electronic noise inside the Drift Tube system using cosmic data. I determined the average noise rate and the temporal and spatial distribution of the noisy cells. Then I analyzed in detail the properties in cosmic data of the DT reconstructed segments as far as their quality, direction, geometrical provenance and resolution are concerned. I devoted a great part of my work in developing (for both the activities described above) a Data Quality Monitoring currently used by the DQM online team for the continuous check of the detector performance status during the data taking.

The CMS muon tracks are obtained combining the trajectories in the muon spectrometer with the ones in the inner tracking system and adding

further information coming from the calorimeter deposits and the muon identification algorithms. These reconstructed muons are then used for the final analyses.

During the last CMS commissioning period, the so-called "Cosmic Run At Four Testla", I studied in detail the muon reconstruction algorithm. Therefore I could isolate the parameters which give the most powerful information on the quality of reconstructed muons and develop the software currently used for the offline monitoring of the muon reconstruction. This software first creates the main distributions of interest and then performs a list of dedicated tests. It eventually gives a definitive report on the muon reconstruction status, declaring the reconstructed muons to be good or not for physics studies.

Finally, I performed a Higgs search analysis using the total amount of data (corresponding to 1 fb^{-1} of integrated luminosity) that will be collected during the first LHC running period at $\sqrt{s}$=10 TeV. I chose a physics process, $H \rightarrow ZZ^{(*)} \rightarrow 4\mu$, with four muons in the final state, to exploit my knowledge on the muon detection and reconstruction in a physics study. The final signal over background significance remains under the threshold of the discovery claim, nevertheless it is a very good result considering the very small Higgs cross section with respect to its main backgrounds and the few statistics available with the first data collected at the LHC. This is due not only to the clear experimental signature of the final state but also to the good muon reconstruction and event selection tuned for this analysis.

Acknowledgement

During my PhD program I was lucky to get in touch with kind and helpful people, fond of their job and able to work together for a common aim. It wouldn't be possible to list all the names, but I want to mention at least the most important to me.

Gianluca, who has always been my reference point on DT detectors; Silvia, for her support on the DT calibration; Riccardo, my main source of knowledge on the CMS muon reconstruction; Sara, for her constant presence and for the help in any question concerning CMS; Giuseppe, for his daily crucial guidance on the DQM framework, and Gianni, who has always assisted me in the CMS control room activities; Nicola, for his aid on my physics analysis planning; the whole group of Torino, for the spontaneity and kindness of all its components. Thanks to all of you!

Many thanks to Chiara, whose suggestions have been fundamental to accomplish this thesis. She has guided me through the Physics associated with the CMS experiment and she has been at disposal for any kind of question concerning my PhD activity.

I would like to address a special thank to Amedeo, who has always shown interest in the progress of my PhD activity. He has given to me many precious advices.

Thanks to all the people who believed in my skills and helped me to find my way for the future.

Finally, I want to say thank you to my flatmates of Saint-Genis: Cristina, Roberto, Daniele, Susy and Marco. Living together has been a unique opportunity of growth in my life. Thanks to Alessandro and Andrea, for their

important support during my stay in Geneve, and thanks to my parents, to Giancarlo, Paolo, Silvia, Cecilia, Vittoria and to all my friends, who are always there for me.

Bibliography

[1] The LHC Study Group, *The Large Hadron Collider Conceptual Design*, CERN/AC 95-05 (1995).

[2] G. Altarelli and M. L. Mangano (editors), *Proceedings of the Workshop on Standard Model Physics (and more) at the LHC* (2000) CERN 2000-004.

[3] T. Sjostrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP 05 (2006) 026. hep-ph/0603175.

[4] H. L. Lai et al., *mproved parton distributions from global analysis of recent deep inelastic scattering and inclusive jet data*, Phys. Rev. **D55** (1997) 1280. hep-ph/9606399.

[5] S. Beole et al., *ALICE technical design report: Detector for high momentum PID*, CERN-LHCC-98-19.

[6] The LHCb Collaboration, *LHCb technical design report: Reoptimized detector design and performance*, CERN-LHCC-2003-030.

[7] W. W. Armstrong et al., *ATLAS: Technical proposal for a general-purpose p p experiment at the Large Hadron Collider at CERN*, CERN-LHCC-94-43.

[8] The CMS Collaboration, *CMS, the Compact Muon Solenoid: Technical proposal*, CERN-LHCC-94-38.

[9] G. Altarelli, M. L. Mangano et al., *Standard model physics (and more) at the LHC. Proceedings*, Workshop, Geneva, Switzerland, May 25-26, October 14-15, 1999, CERN-2000-004.

[10] J. Ellis, *Beyond the Standard Model at the LHC and Beyond*, AIP Conf. Proc. **957** (2007) 38. arXiv:0710.0777.

[11] 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.

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

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

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

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

[16] A. Djouadi, J. Kalinowski and M. Spira, *HDECAY: A program for Higgs boson decays in the standard model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56. hep-ph/9704448.

[17] E. Gabrieli, F. Maltoni, B. Mele, F. Piccinini and R. Pittau, *Higgs Boson Production in Association with a Photon in Vector Boson Fusion at the LHC*, Nucl. Phys. **B781** (2007) 64.

[18] T. Hambye and K. Riesselmann, *Matching conditions and Higgs mass upper bounds revisited*, Phys. Rev. **D55** (1997) 7255. hep-ph/9610272.

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

[20] CMS Collaboration, *The Magnet Project - Technical Design Report*, CERN/LHCC 97-10 (1997).

[21] CMS Collaboration, *The Tracker Project - Technical Design Report*, CERN/LHCC 98-6 (1998).

[22] CMS Collaboration, *The Electromagnetic Calorimeter Project - Technical Design Report*, CERN/LHCC 97-33 (1997).

[23] CMS Collaboration, *The Hadron Calorimeter Project - Technical Design Report*, CERN/LHCC 97-31 (1997).

[24] The CMS Collaboration, *The CMS Physics Technical Design Report, Volume 1*, CERN/LHCC 2006-001, CMS TDR 8.1 (2006).

[25] CMS Collaboration, *The Muon Project - Technical Design Report*, CERN/LHCC 97-32 (1997).

[26] F. Gasparini et al., *Bunch Crossing Identification at LHC Using a Meantimer Technique*, Nucl. Instrum. Meth. **A336** (1993) 91.

[27] The CMS Collaboration, *Calibration of the CMS Drift Tubes Chambers and Measurement of the Drift Velocity with Cosmic Rays*, CMS CFT-09-023, arXiv:0911.4895v1 [physics.ins-det].

[28] J. Puerta-Pelayo, M. C. Fouz and P. Garcia-Abia, *Parametrization of the response of the muon barrel drift tubes*, CERN CMS NOTE 2005/018.

[29] R. Veenhof, *Garfield, a Drift-Chamber Simulation Program user's guide*, CERN Program Library W5050 (1994).

[30] F. Cavallo et al., *Test of MB3 Muon Barrel Drift Chamber with Cosmic Rays*, CERN CMS NOTE 2003/017.

[31] N.Amapane et al., *Offline Calibration Procedure of the Drift Tube Detectors*, CMS NOTE-2007/034

[32] M. Fouz et al., *Measurement of the Drift Velocity in the CMS Barrel Muon Chambers at the CMS Magnet Test Cosmic Challenge*, CMS Note 2008/003 (2008).

[33] The CMS Collaboration, *Performance of the CMS Drift-Tube Local Trigger with Cosmic Rays*, CMS CFT-09-022, arXiv:0911.4893v1 [physics.ins-det].

[34] D.Futyan, ed al., *The CMS Computing, Software and Analysis Challenge*, Proceedings CHEPÕ09, Prague, Czech Republic.

[35] L.Tuura et al., *CMS data quality monitoring: system and experiences*, Proceedings CHEPÕ09, CMS CR- 2009/099.

[36] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987) 444.

[37] C. Liu et al., *Reconstruction of Cosmic and Beam-Halo Muons with the CMS Detector*, Eur. Phys. J. **C56** (2008) 449-460. doi:10.1140/epjc/s10052-008-0674-7.

[38] The CMS Collaboration, *Performance of CMS Muon Reconstruction in Cosmic-Ray Events*, CMS PAPER CFT-09-014, arXiv:0911.4994v1 [physics.ins-det].

[39] C. Mariotti et al., *Search for the Higgs boson in the $ZZ^{(*)}$ decay channel with the CMS experiment*, CMS AN 2008/050

[40] CMSSW Full Simulation Offline Guide, https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideSimulation .

[41] CMSSW Reconstruction Offline Guide, https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideReco .

[42] M. Spira, HDECAY source code for Higgs branching ratios in the SM and MSSM; http://people.web.psi.ch/spira/hdecay/.

[43] C. Amsler *et al.* [Particle Data Group], Phys. Lett. B **667** (2008) 1.

[44] MadGraph/MadEvent v4: The New Web Generation, JHEP0709-028-2007, [arXiv0706.2334].

[45] M. Cacciari, *Updated predictions for the total production cross sections of top and of heavier quark pairs at the Tevatron and at the LHC*, JHEP 0809:127, 2008 [arXiv:0804.2800v3]

[46] J. Campbell and K. Ellis, *MCFM - Monte Carlo for FeMtobarn processes*, http://mcfm.fnal.gov/ *(MCFM version 5.2)*

[47] C. Zecher, T. Matsuura, and J.J. van der Bij, *Leptonic Signals from off-shell Z Boson Pairs at Hadron Colliders*, [arXiv:hep-ph/9404295].

[48] P. Bartalini et al., *NLO vs. LO: kinematical differences for signal and background in the* $H \rightarrow \mathrm{ZZ}^* \rightarrow 4\mu$, CMS NOTE 2006/130 (June 2006) 29 pp.

[49] R.Bellan, Nucl. Phys. Proc. Suppl. **177-178** (2008) 253-254.

[50] T. Lampén et al., *Misalignment Scenarios for the Startup Conditions of the CMS Tracker*, CMS IN-2007/061 (October 2007) 11pp.

[51] S. Abdullin et al., *Sensitivity of the Muon Isolation Cut Efficiency to the Underlying Event Uncertainties*, CMS Note 2006/033 (February 2006) 11pp.

[52] S. Abdullin et al., *Search Strategy for the Standard Model Higgs Boson in the* $H \rightarrow ZZ \rightarrow 4\mu$ *decay channel*, CMS Note 2006/122 (June 2006) 26pp.

[53] S. Baffioni et al., *Discovery Potential for the Standard Model Higgs boson in the* $H \rightarrow ZZ^{(*)} \rightarrow e^+e^-e^+e^-$ *decay channel*, CMS Note 2006/115 (June 2006) 26pp.

[54] S. I. Bityukov, S. E. Erofeeva, N. V. Krasnikov and A. N. Nikitenko, Prepared for PHYSTATO5: Statistical Problems in Particle Physics,

Astrophysics and Cosmology, Oxford, England, United Kingdom, 12-15 Sep 2005.

[55] C. Amsler et al., Physics Letters **B667**, 1 (2008), section 32.3.1

[56] S. Abdullin et al., *Study of pdf and QCD scale uncertainties in pp* $\rightarrow$ $ZZ^{(*)} \rightarrow 4\mu$ *events at the LHC*, CMS Note 2006/068 (April 2006) 10pp.

[57] CMS Collaboration, *Towards a Measurement of the Inclusive* $W\rightarrow \mu\nu$ *and* $Z\rightarrow \mu^+\mu^-$ *Cross Sections in pp Collisions at* $\sqrt{s}$ *=14 TeV*, CMS PAS EWK-2007/002.