Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas

Departamento de Física Atómica, Molecular y Nuclear
Facultad de Ciencias Físicas
Universidad Complutense de Madrid

Measurement of Electroweak processes in muon decay channels, in pp collisions at $\sqrt{s} = 7$ TeV, in the CMS experiment at LHC

Author:
Javier Santaolalla Camino

Supervisors:
Dr. Juan Alcaraz Maestre
Dra. Begoña De la Cruz Martínez
Dra. Isabel Josa Mutuberría

April 11, 2012

A mi querida familia,
especialmente Darío y Aurora
y a *, miembros, socios*
y simpatizantes

Contents

Agradecimientos

Los cuatro años de tesis dan para una larga lista de agradecimientos. Sin embargo, los que me conocen bien saben que, entre lo despistado que soy y lo poco que me gusta dilatarme, muchos agradecimientos no quedarán aqui reflejados, pero también han sido importantes y no quedan en el olvido.

En primer lugar, agradezco enormemente el buen espíritu de trabajo y compañerismo que ha reinado en el grupo del CIEMAT durante los 4 años. Agradecimiento especial a mis tres directores Isabel, Begoña y Juan por lo mucho que he aprendido de ellos y la enorme paciencia que han tenido. Pero también a Marcos, un gran jefe, Jesús, Mara, Juan Pablo, Ignacio... Y por supuesto al resto de estudiantes: Aurelio, Nacho, María, etc. También a mis compañeros de despacho en el CERN (Silvia, Carlo, Óscar, otra vez Juan y Daniel) con los que ha sido divertido trabajar. Sí, divertido. Todos han hecho de los despacho del CIEMAT y del CERN mi segunda, o más bien cuarta o quinta, casa.

Científicamente hablando también debo agradecer mucho a Mary Cruz por la primera parte de la tesis, a María por la paciencia e infinito legado de conocimiento en los W y a Mara por la maravillosa idea de estudiar el W+c. Muy buena idea. Además de todos los coordinadores de los diferentes grupos y todos los que han hecho posible que se generen suficientes W para que yo tenga una tesis. También a la Fundación Mapfre Guanarteme, por ese impulso en el momento más crítico. Mil gracias Julio.

Los que han hecho mis estancias en Madrid y Ginebra más llevaderas, mil gracias. En Madrid, a todo el equipo de fútbol de físicas y amigos de la facultad (Jesús, David, etc.) y a Fete. En Ginebra a los copinos, personal del CERN (María, Natacha, Marc, Nadege and co...), los Zapata invictos hasta hoy, Jojo y el equipo de Challex, Raf, Cari y la salsa y los internos

de UN. Y de manera muy especial a la mejor persona del mundo, Susana. Por supuesto todo no es trabajo, también al sol de Las Palmas, la Playa de las Canteras y todos los que allí me esperaban para desconectar: Néstor y demás noenfermos, Cansino y claretianos y en general a todos los canarios. También a Pedro Hernández y el piano, salvadores más de una vez.

Un agradecimiento muy especial a los que han compartido casa conmigo, pasando momentos inolvidables. Desde los orígenes de LeClub con Albert, y posteriormente los boludos Xabi, Martín, Gastón, Monti, Fran y el profe, y finalmente con Daniel, Garo, Jose y mi primer amigo en el CERN, Tito. A todos los socios, en especial a Carlos, Laurinha, Jabu, Bruno, Byron, Josu y todos los simpatizantes que han hecho grande LeClub.

Burgos, Canarias, Madrid y Ginebra, a todos. Y por supuesto a mi familia, la mejor del mundo. Y en especial gracias a la pequeña Aurora, que no sabe lo que es una tesis pero que tan importante ha sido.

Finalmente, un párrafo especial para una mujer demasiado especial e importante como para compartir párrafo con nadie. Y a toda su familia.

En fin, a todos. A Griffin, Perdue, J.L. Guerra, Cherito, Dieguito, todos los MOWs, Justine, Marylin, Jimena, Dan, Blanka, Ken, Pepe y el Real Madrid, Yayo, Gunther, TitoKing, a todos los cascos azules, Rosana,... GRACIAS.

Viva LeClub.

1

Introduction

November 2009 marked the beginning of the operation era in the Large Hadron Collider (LHC). More than 20 years of conception, design, prototyping, production, test, validation and installation were left behind, and numerous goals successfully accomplished. Four experiments aiming for a wide physical coverage were installed at the LHC: ALICE, ATLAS, CMS and LHCb. The Compact Muon Solenoid (CMS) collaboration chose for the experiment a compact design: a 21 m long, 15 m high cylindrical detector, 12500 tons weight. The design and construction of the different CMS subdetectors were carried out in the approximately 150 institutes of the CMS collaboration.

Proton-proton collisions at the LHC increased steadily in energy and intensity soon after the beginning of the operation. In March 2010 collisions at 3.5 TeV per beam were first ever produced. Operation of the machine during 2010 and 2011 runs has been successful. A dataset with an integrated luminosity of 43 pb^{-1} of data was recorded in the 2010 run, and almost 5 fb^{-1} in the 2011 run, exceeding all the expectations about the LHC operation.

Many physical analyses have been carried out with data from the first two years of operation. Among them, the electroweak measurements stand out. Electroweak bosons are copiously produced in the collisions and their decay in leptons produce a very clear signal for the analysis. Electroweak precision measurements are one of the first to be accomplished.

The importance of the muon particle in this experiment is patent in its own name ("M" of CMS standing for muons). Many of the physical goals of the CMS program are based in searches with muons in the final state. Muon chambers are, indeed, of capital

importance in the CMS design. A fourth of the chambers of one of the technologies adopted as muon detectors in CMS were not only designed, constructed and built by the Spanish CIEMAT institute, but also installed, calibrated and maintained by their technicians and physicists. In November 2009, the performance of the barrel muon chambers was already characterized, being ready for the first LHC collisions and the analysis of thousand of million muons that would arise from the high energy collisions.

This thesis covers different aspects of the CMS operation and data analysis, always with muons as the common factor. The calibration and characterization of the drift tube chambers (DTs), one of the CMS muon detectors, carried out in the pre-collision era, is presented in this thesis. To complete the calibration of the muon detection and measurement, the calibration of muon momentum using electroweak data is shown. In particular, the clean sample that Z bosons coming from collisions represents is used to study the momentum scale and resolution of the CMS muon spectrometer. Finally, electroweak precision measurements are shown. The inclusive electroweak boson production cross section is the first electroweak measurement published by CMS. In this thesis, the inclusive W boson production cross section in the muonic channel with the whole 2010 dataset is measured. As a further step in the study of the W production, this time accompanied by jets, the analysis of the exclusive final state W+charm jets is both challenging and interesting from the physical point of view. The separation of W+charm from other W+quark jets is successfully achieved using c-tagging techniques that are new in CMS and useful for any other c jet studies. This analysis can constrain the quark strange content in the proton. Prospects to improve the measurement with 2011 data, as well as important distributions with already 2 fb^{-1} are also reviewed in the last part of the thesis.

The description of the LHC and its physical goals are shown in Chapter 2. The CMS detector is also presented in detail in this Chapter.

Chapter 3 is devoted to muon chambers. The importance of muons in the CMS physics program, the design of CMS DT chambers and the characterization and calibration of this subdetector with cosmic muons in the pre-collision era are shown. The fundaments of the detection using drift tube chambers are discussed. The calibration procedure is revised and completed with an analysis based on the drift velocity measurement and the factors that can modify the result of this measurement. The DT chambers are characterized in terms of reconstruction efficiency in the same Chapter.

The proposed method is applied to scenarios with the solenoidal magnetic field on and off to compare the results.

The muon transverse momentum calibration is the topic of Chapter 4. The method presented aims to compute the extra momentum scale and resolution terms to complete the already well performing simulation of physical processes through CMS and adequate it to the experimental data. Many analysis can benefit from this method, since it directly provides the distortion factors that can be used to make simulated samples resemble data. The application of this method to evaluate systematic uncertainties based on muon momentum scale and resolution is straight forward. This "bridge" Chapter uses elements of the first part of the thesis (muon detection) and introduces elements of the second part of the thesis (electroweak boson selection with collision data).

Chapter 5 aims to underline the importance of electroweak studies in any collision experiment, to set the basis for W boson selection and to compute the W boson production inclusive cross section. This analysis benefits from high statistics recorded in the 2010 run (already dominated by the systematic uncertainties). This allows to present a precision measurement already at the 2% level that is shown to be in agreement with the Standard Model. The use of templates based on experimental data to model shapes of distributions for signal and background processes, the efficiency computations and other techniques that exploit the huge amount of electroweak boson recorded are a key part to reduce these measurement uncertainties to the minimum achieved so far in CMS.

Chapter 6 is devoted to the associated charm production in W boson final states. Its application in terms of parton densities for the strange quark and the implication that an improvement on this side could have in other precision measurements are presented to motivate this study. A detailed selection and extraction of the W+c signal is discussed in this Chapter, leading to a clear signal signature by using c-tagging techniques. The ratios, both of positive over negative W+charm production and the W+charm over the W+jets, are presented and the uncertainties in this measurement shown in detail. Prospects to exploit the possibilities to reduce the systematic uncertainties with larger data samples and the sensitivity of this study to fulfill PDF groups requirements on the strange quark composition of the proton are also discussed.

2

The Large Hadron Collider and the CMS experiment

The LHC (Large Hadron Collider) is a 27 km synchrotron that produces proton-proton collisions at a new energy scale (1, 2, 3). It is the largest and most powerful particle accelerator at present. The tunnel that hosts the accelerator is located near Geneva (Switzerland), between 50 and 125 meters under the surface. During the first two years of operation, the LHC has produced collisions at an energy of 7 TeV (center-of-mass system). After these two years of operation, the amount of data recorded is enough to perform high precision physical measurements.

2.1 LHC physics program

The current model for particle physics is called Standard Model. It is accurate and almost complete (4). The SM embodies our current understanding of particles and forces. It was developed in the middle 20th century. This theory has been tested with an extraordinary precision in many different experiments. It has explained a great deal of experimental results and predicted a host of varied phenomena. Today this model has become established as a well tested physical theory and a reference model to describe most of the phenomena.

The SM has successfully described hundreds of particles and their complex interactions with few ingredients: 12 matter particles and their corresponding antiparticles (see Table 2.1); and force carrier particles (see Table 2.2). The matter particles (fermions)

	Charge	First Generation	Second Generation	Third Generation
Leptons	-1	Electron (e)	Muon (μ)	Tau (τ)
	0	Neutrino(ν_e)	Neutrino (ν_μ)	Neutrino (ν_τ)
Quarks	-1/3	Down (d)	Strange (s)	Bottom (b)
	2/3	Up (u)	Charm (c)	Top (t)

Table 2.1: Matter content as described in the SM (5).

Electromagnetic	Weak	Strong	Gravitation
Photon (γ)	Weak Bosons (W^+, W^-, Z^0)	Gluons (g)	Graviton (g) ?
m = 0	m $\sim$ 100 GeV	m = 0	m = 0

Table 2.2: Force carrier particles as described by the SM.

are classified as quarks or leptons depending of whether or not they feel the strong force respectively. The interactions in this model are described by the interchange of a force carrier particle (bosons). The SM has successfully integrated three of the forces known in nature: electromagnetic, weak and strong. Forces in this theory are described as quantum fields in which the Lagrangian is invariant under a continuous group of local transformations. Quantum electrodynamics, the theory describing the electromagnetism within this model, is described using an abelian gauge theory with the symmetry group U(1) with one gauge field, and the photon as the gauge boson. By adding to this theory the symmetry group SU(2) the SM has been able to describe also the weak force, reaching the unification of both forces, that are shown to be different aspects of the same force. At the energy of the order of 100 GeV both forces appear to be the same. The force carrier particles of this unified gauge group U(1)×SU(2), after the symmetry breaking, are called photon, W^+, W^- and Z.

The strong force is included in this model using a similar description: a quantum field with a symmetry group. In this case SU(3) is the non-abelian group that successfully describes the strong processes. This quantum theory is called quantum chromodynamics (QCD) and is added to the SM by increasing the gauge group to form the final U(1)×SU(2)×SU(3).

In the U(1)×SU(2) Lagrangian the force carrier particles appear to be massles. However, experimentally the masses of the W and Z bosons result to be different from 0. The electroweak symmetry breaking mechanism was introduced in this model as a

solution to this problem. The Higgs mechanism, implemented by a Higgs field and its interaction with the other fields, is the missing element within the theory in order to produce the mass terms. The way these bosons acquired mass explained at the same time the masses of all particles within the SM. It is the simplest mechanism capable of giving mass to the gauge bosons while remaining compatible with gauge theories. The Higgs field would consist of 4 component fields, two charged and two neutral. Three of them are used to give mass to the electroweak bosons (W^+, W^- and Z). Both charged scalar and one neutral components are goldstone bosons which act as the longitudinal third polarization components of the massive W^+, W^- and Z bosons. The remaining degree of freedom would become detectable through a boson, the Higgs boson: a 0-spin massive particle. All the properties of this particle, except the mass, are predicted by the theory.

Despite the accuracy of the SM describing many of the physical phenomena, our understanding of the Universe is still incomplete. In order for the SM to be a complete theory, the Higgs boson has to be discovered. Direct searches in the former CERN main accelerator (LEP) excluded the existence of such a particle if its mass is smaller than 114 GeV (6, 7). Measurements carried out in the proton-antiproton collider at Fermilab (Tevatron) excluded the existence of this boson in the mass window [158 - 176] GeV (8). More recently, in the LHC, the ATLAS and CMS experiments combined exclusion range for 2.3 fb^{-1} of proton-proton collision analyzed spans from 141 to 476 GeV (9). In Fig. 2.1 the exclusion region for CMS and ATLAS (combined), LEP and Tevatron is shown. In Fig. 2.2 the preferred Higgs boson mass, following a fit to the measured parameters of the SM is shown. From both figures we can conclude that even if the Higgs boson is excluded for most of the window mass where it could exist, the preferred region needs still some time to be analyzed. In the short future, due to the high luminosity-high energy collisions at LHC, the experiments will cover the full theoretically permitted mass spectrum for the Higgs boson. At LHC both ATLAS and CMS experiments were specially designed to optimally search for this missing particle.

If the Higgs boson is proved not to exist, many of the different theoretical proposals based on a higgless scenario should be tested experimentally. LHC may shed light over all these possibilities. The SM, despite being a successful theory to describe many physical phenomena, is today considered a phenomenological theory part of a more fundamental one yet to be described. Among SM limitations, the most important are:

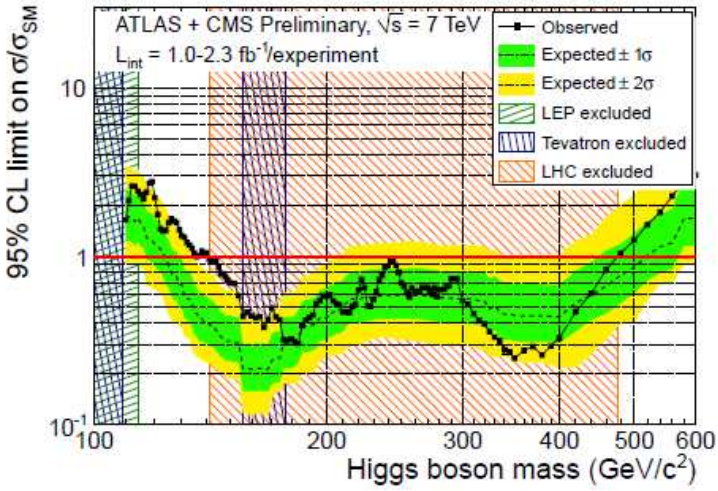

Figure 2.1: Exclusion region (LEP, TEVATRON and the combination of ATLAS and CMS) - Excluded region with 2.3 fb^{-1} of proton-proton collision available. Exclusion bands coming from LEP and TEVATRON are also shown.

- it does not include gravity. Gravitation is the only known force which can not be described using SM formulation: quantum fields with gauge symmetry. Gravity is accurately formulated using Einstein's Relativity, which is a relativistic, non-quantum theory. Every attempt as of today to quantize gravitation has not been successful. The description of all physical phenomena within a unique framework is the ultimate goal of physics, called unification theory. Such kind of theories, globally called GUT, have been actively looked for during the last decades. The LHC by exploring a new energy range will provide useful information in the chase of this theory.

- it has too many parameters. The inclusion of many parameters in the theory makes the model less predictable and elegant, without an apparent reason behind each particular value. The existence of so many parameters in the SM make the physicists community consider SM not to be the fundamental theory of nature.

- even if the missing piece of the SM is discovered in the LHC, there is a problem that will remain unsolved. Unless there is a fine-tuning cancellation in loop diagrams, the large quantum contributions to the square of the Higgs boson mass would make this to be bigger than its expected value. This is the so called

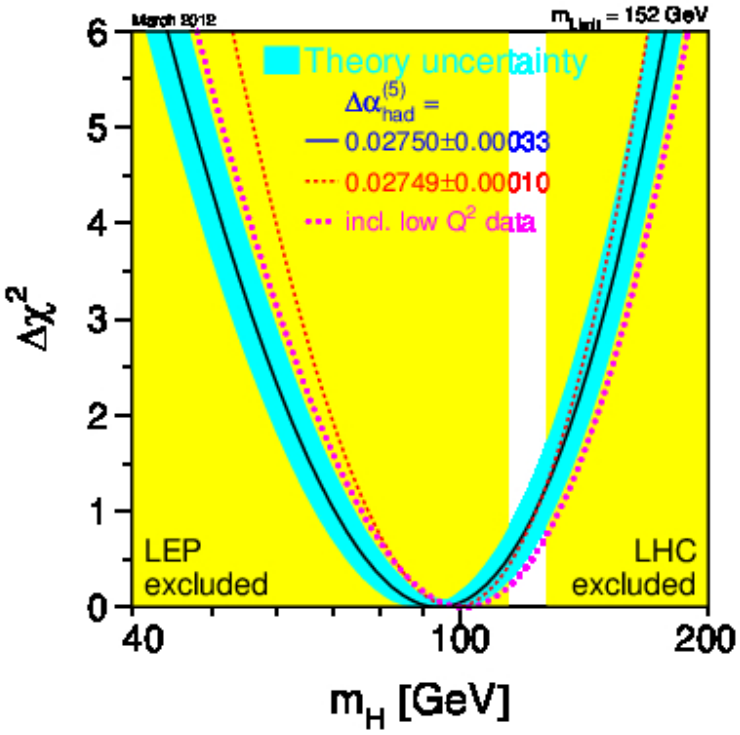

Figure 2.2: Exclusion region - The χ^2 for a global fit to electroweak data is shown as a function of the Higgs mass. The solid line corresponds to the result of this fit with the blue band indicating the impact of theoretical uncertainty. The yellow-shaded area indicates the region of Higgs masses excluded by experiment.

hierarchy problem. Cancellations provided by a new group of particles is the solution proposed by Supersymmetry to solve this problem.

- there is no energy scale at which electromagnetic, weak and strong coupling constants converge in a unique value. Again, Supersymmetry is a proposed theory which would solve this problem.

Many different theories were proposed during the last half century to substitute the SM. Among them supersymmetry is the most popular (10). Sypersymmetry is an extension of the SM. By doubling the number of particles it is able to solve the hierarchy problem, and, at the same time, unify the three described forces. It adds a new symmetry to those of SM: fermion-boson symmetry. In this theory, the number of bosonic number of degrees of freedom equals the fermionic ones. As a consequence, integer spin particles with the same properties as leptons and quarks should appear: the so called sparticles. If this theory is correct, the sparticles may appear in the proton-proton collisions at the LHC. Other theories such as string theory (11), extra dimensions or technicolor (12) are also expecting new LHC data to be confirmed or rejected.

Following cosmological observations, the amount of matter in the universe described by SM is only 4%. The other 96% remains unknown, and it is shared between two different kinds of energy/matter whose origin is unknown, called dark matter (23%) (13) and dark energy (73%). The existence of the dark matter was proposed after several indirect cosmological measurements. So far, no hints of any dark matter candidate have appeared in any of the different experiments carried out for such a purpose. At LHC these particles may be created and studied in the two major experiments (CMS and ATLAS).

Another open question in particle physics is the matter-antimatter asymmetry. If matter interacts exactly in the same way antimatter does, only radiation would have remained after matter annihilation at the Big Bang. Due to the fact that there is matter in the Universe and matter and antimatter are not coexisting, there should be a difference in their interaction, namely, that there exists a violation of the CP (charge-parity operator) conservation (14). The solution to this problem is also one of LHC goals, and a dedicated experiment, LHCb, has been setup for it.

Besides new discoveries of particles in the collisions, the physics program at the LHC is wider, including studies in many fields of particle physics. Electroweak studies are important for several reasons. First of all, they provide a test to the validity of the SM at a new energy regime. The improvement on the knowledge of some Electroweak boson properties, e.g. W boson mass, is useful to better understand and constrain the SM (15). Several Electroweak measurements and production channels can improve the knowledge on the proton parton composition (PDFs). Electroweak processes are the main background of many discovery channels, and their understanding directly affects these searches. They also provide high mass resonances whose decay products are used to calibrate the detector response.

Due to the available energy in the collisions and the dominant gluon-gluon fusion production process, the LHC provides a huge amount of top-antitop events, making LHC a top quark factory. Top quark was the last particle of the SM to be discovered (16). By measuring its properties, we are sensitive to new physic phenomena. Top quark production is also a main background for many interesting processes and a tool for calibration and physics commissioning purposes.

The study of medium-mass-quark production in high-energy hadronic interactions plays a critical role in testing next-to-leading order (NLO) Quantum Chromodynamics (QCD) calculations. Additionally, the hadroproduction of quarkonia is not fully understood. For example, none of the existing theoretical models satisfactorily describes the prompt J/ψ differential cross section. LHC results are improving our understanding of the quarkonium productions mechanisms. Low mass resonances like the J/ψ also contribute to the calibration of the detector.

2.2 The LHC

The LHC is the most energetic particle accelerator built up to now (1, 2, 3). It collides bunches of protons at almost the speed of light. Before acceleration in LHC tunnel, from 450 GeV till nominal energy (3.5 TeV), protons are pre-accelerated using different CERN machines (see Fig. 2.3). Protons are extracted from hydrogen atoms. The first accelerator is linear, Linac2, in which protons reach an energy of 50 MeV before being injected in the following step: the PS Booster. The PS Booster is a synchrotron

accelerator of 25m radius which accelerates protons to feed the next step in the accelerator sequence: the Proton Synchrotron (PS). Protons reach PS with an energy of 1.4 GeV. PS is a 628 m circumference ring which accelerate protons that are injected at an energy of 25 GeV into the super Proton Synchrotron (SPS). The SPS accelerates them to reach an energy of 450 GeV per bunch and transfers them to the LHC.

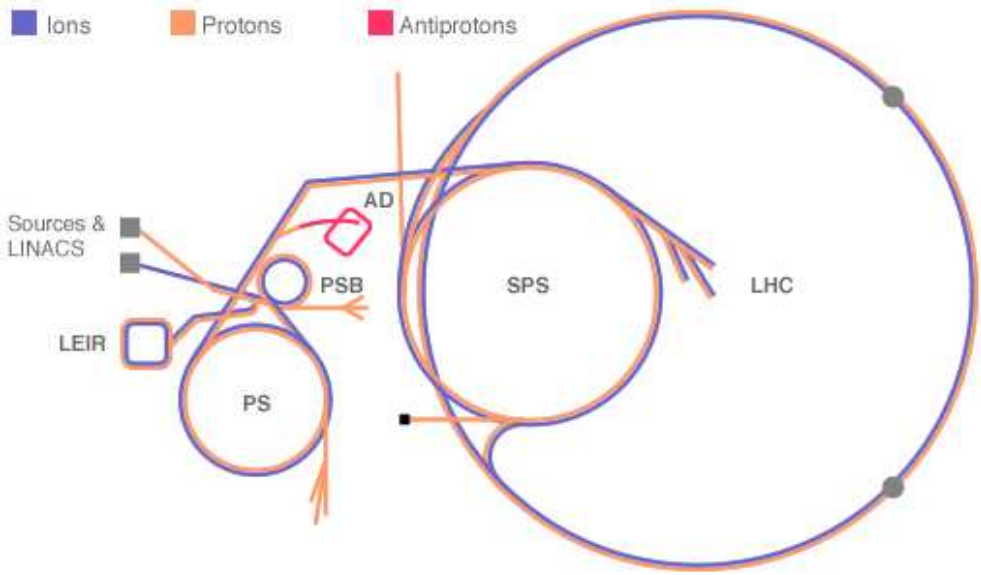

Figure 2.3: CERN accelerator complex - LHC and all the accelerators involved in the experimental complex

The LHC is hosted in the LEP tunnel, a 27 km long circumference, with 25% of its perimeter in Switzerland, the rest under French surface. The tunnel was built 100 m underground in average. At the same time, Earth provides a good shielding for radiation. It was built with a 1.4% tilt making its depth vary between 175m and 50m.

LHC accelerator is made of arcs and insertions. The arcs contain each 154 bending superconducting magnets. Insertions are straight sections plus two transition regions placed in each end. Depending on the use of the insertion, its layout could vary. There are insertions for physics, injection, beam dumping and beam cleaning. A sector is a segment of LHC with an arc and an insertion. LHC consist of 8 sectors, which are defined as the accelerator segment between two insertion points.

The high mass of the particles of interest makes necessary collisions with particles at high energy. The small probability for these particles to be created with respect

to other non-physically-interesting processes makes necessary a high collision rate and a small interaction area. In other words we need collisions at high energy and high luminosity. These are the two most important parameters for an accelerator.

To reach this high energy, it was necessary to build a long accelerator ring (27 km) with high magnetic field to bend the proton beam (8.3 T). To accelerate the particles, superconducting radiofrequency cavities are used. Currently the protons are accelerated to an energy of 3.5 TeV. In the near future (2014 or 2015) it is foreseen to reach 7 TeV, at full magnets power.

Luminosity depends on beam parameters (2). For a gaussian beam profile distribution:

$$L = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F \tag{2.1}$$

where N_b is the number of particles per bunch (10^{11} protons at LHC), n_b the number of bunches per beam, f_{rev} the revolution frequency (approx. they turn 11 thousand times per second), γ_r the relativistic gamma factor, ϵ_n the normalized transverse beam emittances, β^* the beta function at the collision point and F the geometric luminosity reduction factor due to the crossing angle at the interaction point. Given the number of particles per bunch and number of bunches in the beam, the luminosity can be increased by reducing the emittances and the β factor. LHC uses corrector magnets to avoid the dispersion of the beam and thus increase the luminosity. By the end of September 2011, after two years of proton-proton collisions, LHC reached a peak luminosity of $3 \cdot 10^{33} cm^{-2} s^{-1}$ and it has kept working at this value until the end of 2011 proton-proton data taking, in November 2011.

These high values made the design of the LHC a technological challenge, implying development of new technologies in different fields:

- Vacuum. Protons in the accelerator must travel in the beam pipe in the absence of matter to avoid scattering or energy loses. LHC has three vacuum systems: for cryomagnets, to avoid heat load by gas conduction, it requires a pressure of only about 10^{-6} mbar; for the helium distribution line and for the beam vacuum (reaching 10^{-9} mbar).

- Magnets. There are about 9600 magnets in the LHC, 1232 of them are 15 m long, 35 tons dipoles. Dipoles use NbTi cables working at 1.9 K. The magnet coils

for the LHC are wound from 'Rutherford' cable. This cable consists of up to 36 twisted 15-mm strands, each strand being made up in turn of up to 8800 individual filaments, each filament having a diameter as small as 7 micrometres. At this temperature these cables conduct electricity without resistance (superconductor). The current, 11850 A, is enough to create the 8.3 T magnetic field required by design. The other magnets (quadrupoles, octupoles, decapoles,...) are used to focus and squeeze the beam and optimize its trajectory. Most of them are embedded in the main dipoles or quadrupoles.

- Cryogenics. The LHC is the biggest cryogenic system in the world. To maintain its 27 kilometre ring at superfluid helium temperatures, the LHC's cryogenic system have to supply an unprecedented total refrigeration capacity of some 150 kW at 4.5 K and 20 kW at 1.9 K distributed around the ring.

- Cavities. They are aimed to deliver radiofrequency power to the beam during energy ramping. They also keep the proton bunches in LHC as tight as possible in order to reduce the bunch spread, increasing as a consequence the instant luminosity. LHC uses superconducting cavities due to their small losses and large stored energy. In total, 8 cavities per beam are used, placed in a long straight section.

Collisions are designed to take place at LHC in the 4 interaction points, where the tunnel turn into a cavern to albergate the detectors. The 4 main LHC experiments are placed in these caverns: ALICE (A Large Ion Collider Experiment) (17), ATLAS (A Toroidal LHC AparatuS) (18), CMS (Compact Muon Solenoid) (19) and LHCb (LHC experiment for b-physics) (20). TOTEM and LHCf, two forward smaller experiments, are placed nearby the interaction point of the two big experiments at LHC (CMS and ATLAS) and aimed to study very forward physics processes.

ATLAS and CMS are general purpose detectors, they aim to look for any hint of new physics in the new available energy spectrum. LHCb will study the physics of b quarks (trying to solve the matter-antimatter mystery). ALICE will study heavy ion collisions, trying to observe a new state of matter called quark-gluon plasma.

LHC operation started the 23th November 2009 with the first proton-proton collisions. In these collisions, the proton beam consisted in only one bunch of protons and

collided at the injection energy (450 GeV per proton). Shortly after, the 30th March 2010, the beam energy was increased to an energy of 3.5 GeV per proton, starting the so called 2010 run. It finished the 8th November 2010 with a peak luminosity of $2 \cdot 10^{32}$ $cm^{-2}s^{-1}$. During this period, an integrated luminosity of 36 pb^{-1} of proton-proton collisions at 3.5 GeV per proton were collected and certified as good data for analysis in CMS and ATLAS.

The 2011 data taking period started on the 13th of March of 2011 with collisions at 3.5 GeV per proton beam at a peak luminosity of $1.2 \cdot 10^{30}$ $cm^{-2}s^{-1}$. The last run considered in this thesis involved collisions at 3.5 GeV of beam energy and was taken the 7th of September 2011 at an instantaneous luminosity of $3 \cdot 10^{33}$ $cm^{-2}s^{-1}$. This period (2011 runA) was followed by a technical stop aimed for improvement of the accelerator. The total luminosity collected and certified by CMS during this period is 2.3 fb^{-1}.

2.3 CMS

CMS is a cylindrical detector which is placed in one of the LHC interaction points. It is a 21.6 m length, 14 m diameter cylinder of 12500 tons, with a solenoidal field of 3.8 T (21) (see Fig. 2.4). This design guarantees a full coverage around the interaction point and a good transverse momentum resolution, for an accurate measurement of the physical processes generated at the collisions.

Being a general purpose experiment, it explores any hint of new physics at the LHC, besides performing precision measurements on existing particles. The goals pursued by this experiment determine its main characteristics (21, 22):

- Muons. Redundant and accurate muon system, with excellent muon identification, over a wide range of momenta in the pseudorapidity region below 2.5. Good dimuon mass resolution and negligible charge miss-identification for muons with momentum lower than 1 TeV/c. Excellent trigger efficiency. The importance of muons in CMS will be seen in Chapter 3.

- Tracker (pixel and silicon strips). Good performance in the tracker: excellent momentum resolution and reconstruction efficiency. High trigger efficiency and b-jet tagging efficiency. CMS tracker is the biggest silicon detector ever built.

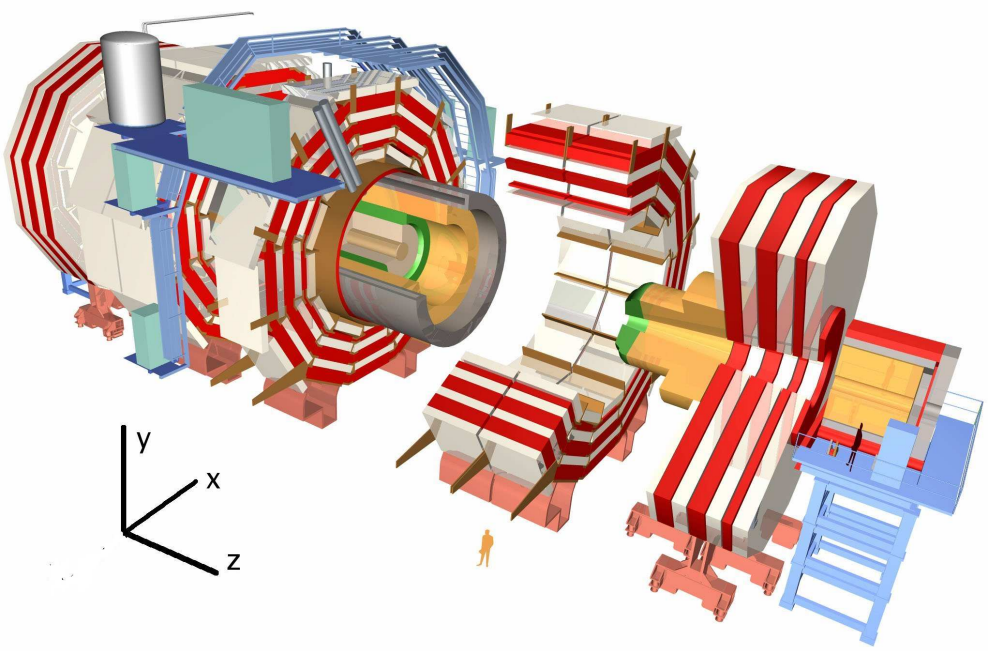

Figure 2.4: CMS detector - Representation of CMS detector

- Electromagnetic calorimeter. Exceptional energy resolution, diphoton and dielectron mass resolution. Wide geometric coverage, good direction measurement, π^0 rejection and efficient photon and lepton isolation. Its design was driven by the needs to cover the low mass ($\approx$ 100 GeV) region for the Higgs boson ($H \rightarrow \gamma\gamma$ channel).

- Hadronic calorimeter. Reasonable dijet mass resolution and missing transverse energy resolution. Large coverage and fine lateral segmentation.

In CMS different detector technologies are applied in order to optimize particle identification. These detectors are placed concentrically, and in two regions: barrel (central part of CMS covering till 1.1 in pseudorapidity) and endcaps. In the central region (the barrel, see Fig. 2.5), going from the innermost part to the outermost one we find: a tracker detector, built with silicon strips and pixels, the electromagnetic calorimeter (ECAL), the hadronic calorimeter (HCAL) and the muon chambers. In the barrel region drift tube chambers and resistive plate chambers are used as muon detectors, and are located in 5 wheels. In the 2 endcaps, cathode strip chambers are

installed instead due to occupancy and magnetic field conditions in this part of CMS. A brief description of each of the subdetectors follow.

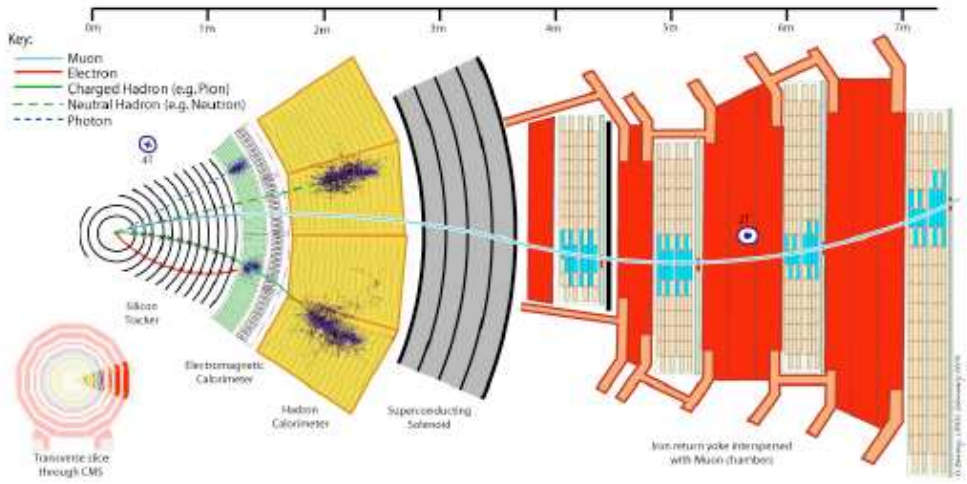

Figure 2.5: CMS detector transversal slice - Slice of CMS detector showing the different technologies used in CMS barrel.

- **Magnet**. The aim of CMS magnet is bending the charged particles in the transverse plane to compute, by means of the radius of curvature, the momentum of the particle in this plane. CMS magnet is a 13 m long, 5.9 m diameter superconducting solenoid producing a 3.8 T magnetic field. The muon chambers in the barrel, in the outer side of the solenoid, benefit from a high return field. They are placed in the 1000 ton return yoke, in the barrel region. A 2 T residual field in this region allows a good momentum resolution in the muon chambers. The stability of the magnetic field in this outer region is also an asset, to ease the muon reconstruction and to allow a good momentum resolution (23). The calorimeters and tracking detectors are placed within the solenoid. CMS magnet consist of 5 modules. Each module contains 4 layers winding of NbTi cables. With this design (19.5 kAmperes current, 2.6 GJ stored energy) the quality factor (stored energy over mass) is improved by a factor two with respect to the previous state-of-art magnet designs.

- **Inner Detectors**. They are built to accurately measure the trajectory of the charged particles, in a region where the magnetic field is intense. They aim

to compute the momentum of the particles in the transverse plane, to detect the vertices and for triggering purposes. CMS inner detector consist of silicon detectors (24, 25). It uses three different designs depending on the expected flux in the three different regions (see Tab. 2.3). In the innermost part, close to the interaction point, where the flux is very high ($10 < r < 20$ cm) pixel detectors are used. Given the size of each pixel (100×150 μm^2) and the luminosity in nominal operation mode, the occupancy per channel for this subdetector is about 10^{-4}. In the intermediate region ($20 < r < 50$ cm) CMS makes use of silicon microstrips. The expected rate of particles in this region allows the use of 10 cm $\times$ 80 μm detectors, keeping the occupancy low enough (2-3%). Larger-pitch silicon microstrip are used in the outermost region of the inner detector, where the flux of particles has substantially decreased ($r > 55$ cm). The maximum cell size (25 cm $\times$ 180 μm) is still small enough to keep occupancy below 1%.

Detector	Radius	Size	Occupancy
Pixels	$r > 10$ cm	100×150 μm^2	$< 10^{-4}$
Silicon microstrip	$20 < r < 50$ cm	10 cm $\times$ 80 μm	2-3%
Silicon microstrip	$r > 55$ cm	$<$ 10 cm $\times$ 80 μm	1%

Table 2.3: Description of the different technologies and sizes of the tracker detector components in CMS.

Pixel detector. The total amount of pixels in CMS is 66 millions, covering an area of 1m^2. They are distributed in two areas: barrel and endcap (see Fig. 2.6). In the barrel, 3 pixel layers are placed at 4.4 cm, 7.3 cm and 10.2 cm of average radius. 768 pixel modules, arranged into half-ladders of 4 modules each are used in this region. The total length of the pixel detector in this part is 53 cm. In each of the 2 endcaps CMS uses two pixel disks. The detector in this region extends from 6 to 15 cm in radius, with 672 pixel modules, with 7 different modules in each blade. The spatial resolution obtained is 10 μm in r-ϕ plane (transverse plane), 15-20 μm in the z coordinate (longitudinal axis).

Strip detector. It consist of 9.6 million silicon strips located in 15400 modules, covering an area of 200 m^2. Strips in CMS are placed in 2 regions in the barrel (TIB, Tracker Inner Barrel, and TOB, Tracker Outer Barrel) and two regions in

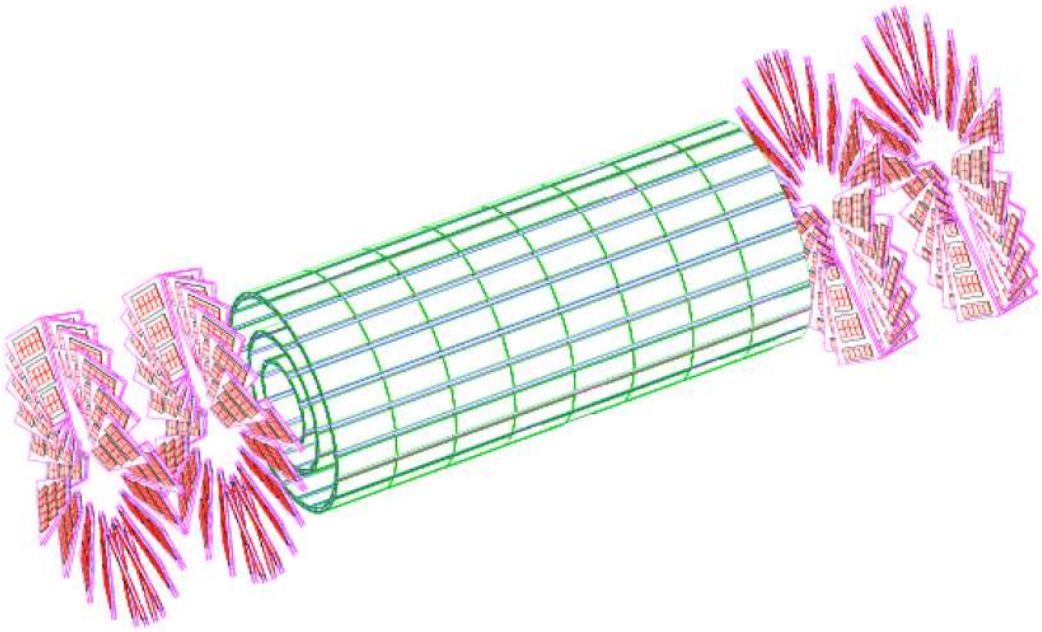

Figure 2.6: Layout of the pixel detector - Layout of CMS pixel detector. It is divided in two regions: barrel and endcap

each of the endcaps (TID, Tracker Inner Disks and TEC, Tracker End Cap) (see Fig. 2.7 and 2.8). In the barrel region, the tracker detector of CMS consists of 10 layers: 4 layers in the TIB, and 6 in the TOB. The TIB covers up to $z < 65$ cm. Strips in this region are 320 μm thick, and 80-120 μm pitch. Resolution in the TIB is 24 μm approximately in both r-ϕ plane and z direction. In TOB thicker (500 μm) and wider pitch (120-180 μm) detectors are used, since the flux in this region is lower. Resolution in r-ϕ plane varies from 32-52 μm, and in z direction 52 μm. Each TEC consists of 9 disks. They are arranged in rings, having different pitches. The thickness is 320 μm in the innermost regions and 500 μm for the rest of the TEC. TID comprises 3 small disks. Strips in this region are 320 μm thick. The performance of the tracker detector is shown in Fig. 2.9 and Fig. 2.10. In Fig. 2.9 the resolution computed with first LHC data is shown. The resolution is computed after a fit to the mass of dimuon events candidates to come from J/ψ resonance, versus the muon (lepton) pseudorapidity. A $\approx 1\%$ resolution is observed in the central region whereas a $\approx 3\%$ resolution is computed for the endcaps. In Fig. 2.10 the primary vertex resolution is plotted versus the number

of tracks for different average p_T and in the 3 spatial axis. When the number of tracks used to determine the primary vertex is sufficiently high, the resolution achieved in the transversal plane is of the order of 10 μm and $\approx 15 - 20 \mu m$ in the longitudinal one.

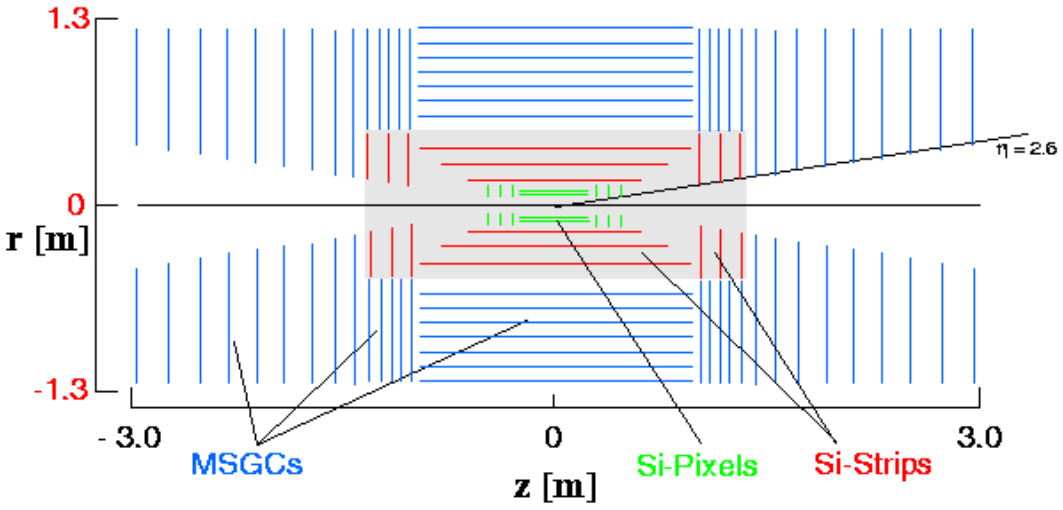

Figure 2.7: Layout of the tracker detector - Layout of CMS tracker detector.

- **Electromagnetic Calorimeter (ECAL)**. It is essential in the identification of electrons and pions and in the measurement of the energy in hadronic showers (21, 28). This detector is placed inside the solenoid, covering a pseudorapidity range of $|\eta| < 3.0$. Lead tungstate ($PbWO_4$) crystals are used for being fast (80% of the light is emitted within 25 ns), compact, and their fine granularity and radiation resistant, having an excellent energy resolution. However, these crystals present the drawback of a relatively low light yield (30 γ/MeV). Photodetectors with intrinsic gain, and that can operate in a magnetic field, are used to detect the radiation coming from the electromagnetic cascade and its amplification. Silicon Avalanche Photodiodes (APDs) are used in the barrel (EB), Vacuum Phototriodes (VPTs) in the endcaps (EE). In total, 61200 lead tungstate crystals in the barrel part and 7324 in the 2 endcaps comprise CMS ECAL. Details on the EB and EE elements are given in Tab. 2.4.

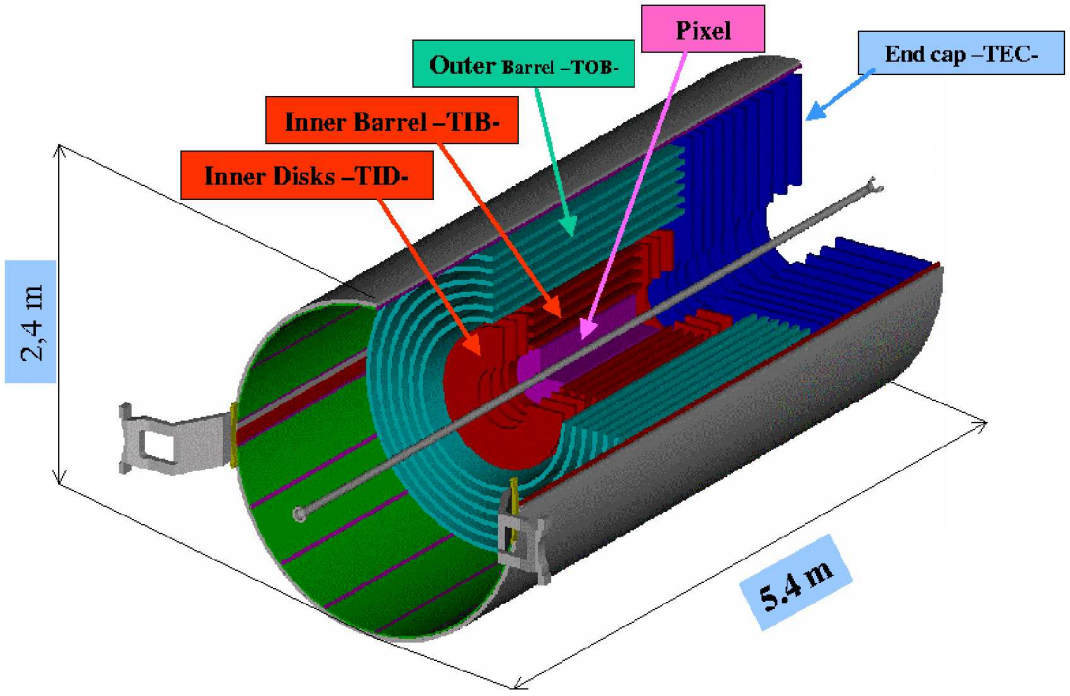

Figure 2.8: Representation of CMS tracker detector - Pictorial representation of CMS tracker detector.

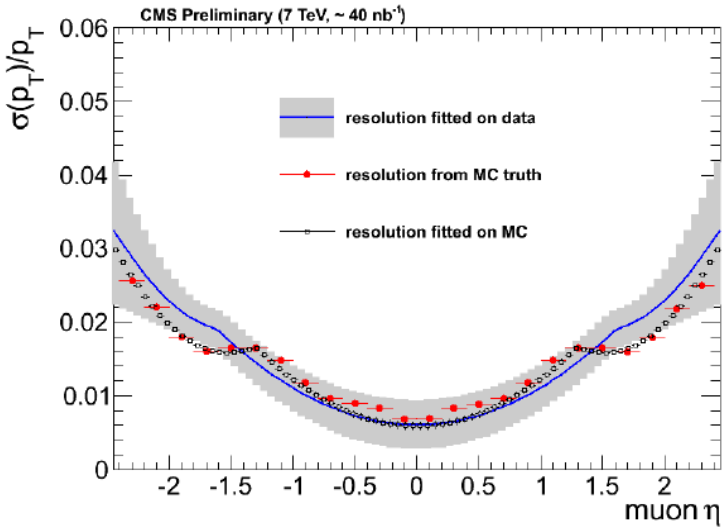

Figure 2.9: Resolution on transverse momentum. - Resolution on transverse momentum with data (blue line) compared to the MC resolution computed from MC truth (red points) and from the fit using dimuon events from the decay of the J/ψ resonance (26).

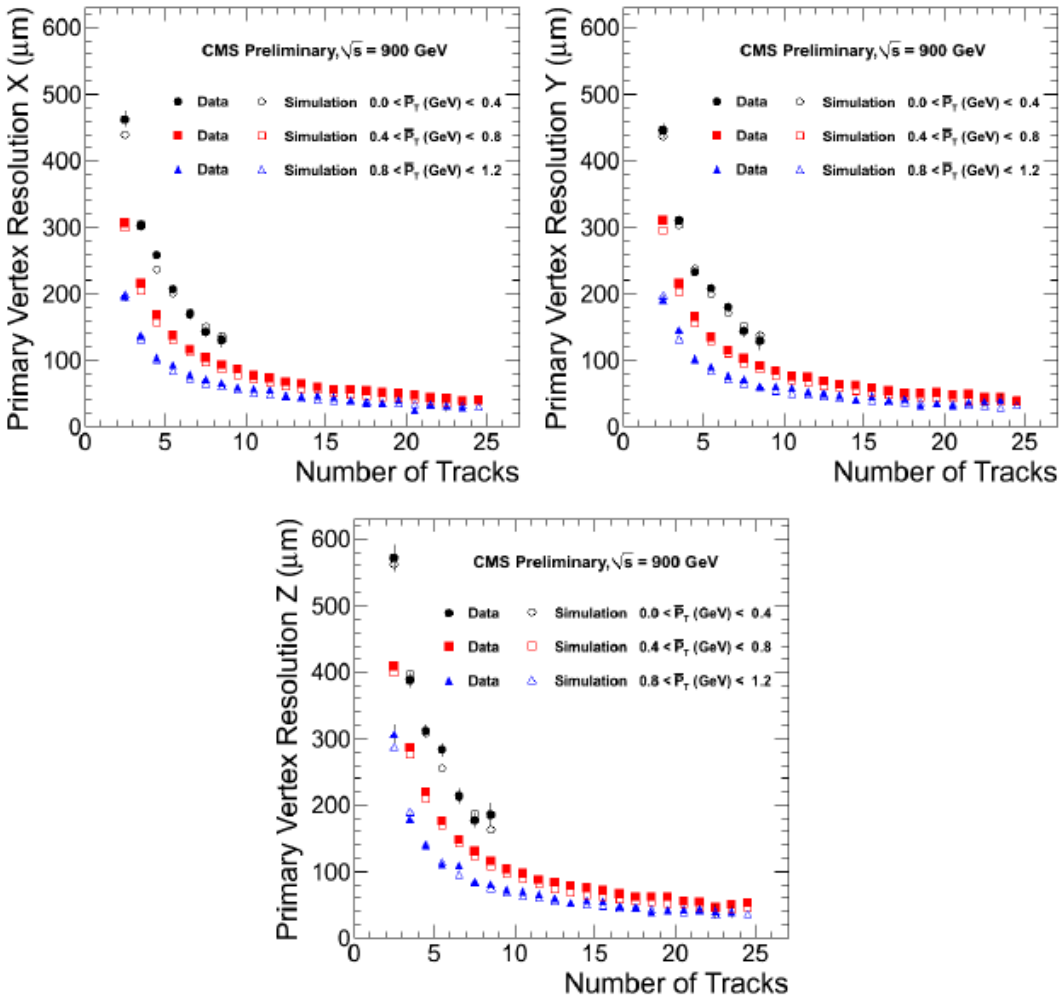

Figure 2.10: Vertex resolution - Primary vertex resolution in the X axis (top left), Y axis (top right) and Z axis (bottom) versus the number of tracks for different average p_T (27).

In the barrel (EB) 22×22 mm^2 front face cross-section crystals are used. Their length (230 mm) corresponds to 25.8 radiation lengths. EB has an inner radius of 129 cm and consist of 36 identical modules, each containing 1700 crystals. It covers a pseudorapidity range $|\eta| < 1.479$. The granularity obtained is of $\delta\eta \times \delta\phi = 0.0175 \times 0.0175$.

The endcaps (EE) use 28.6×28.6 mm^2 front face cross-section crystals, 220 mm length (24.7 radiation lengths). They are placed 314 cm away from the interaction point, covering a η range of $1.479 < |\eta| < 3.0$. Crystal are grouped in "supercrystals", units of 5×5 crystals. A preshower device, 2 planes of silicon strip detectors with disks of lead absorber, is placed in front of the crystal calorimeter to improve the position determination of the particles, and pion identification.

	Cross section	Length (Radiation Lenghts)	Coverage (η)
EB	22×22 mm^2	230 mm (25.8)	$\|\eta\| < 1.479$
EE	28.6×28.6 mm^2	220 mm (24.7)	$1.479 < \|\eta\| < 3.0$

Table 2.4: Description of the elements of the ECAL detector.

In Fig. 2.11, the energy resolution for electrons from simulation as a function of the incident particle energy is shown. This curve is parametrized as a function of energy for energies below 500 GeV, by fitting a gaussian function to the reconstructed energy distribution:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \tag{2.2}$$

where S is the stochastic term, N the noise and C the constant term.

In Fig. 2.12 the energy spectra of the ECAL channels in the barrel and the endcaps in proton-proton collisions are shown. In Fig. 2.13 the observation of π^0 when decaying to two photons coming from proton-proton collisions is shown.

- **HCAL**. The aim of the hadronic calorimeter is to measure the energy of hadron showers (21, 30). Its requirements are mostly imposed by an accurate measurement of the missing transverse energy (MET). As a result, hermeticity is an asset. Minimizing the non-gaussian tails in the energy resolution also imposes stringent requirements in the HCAL design. The absorption capacity is another constraint,

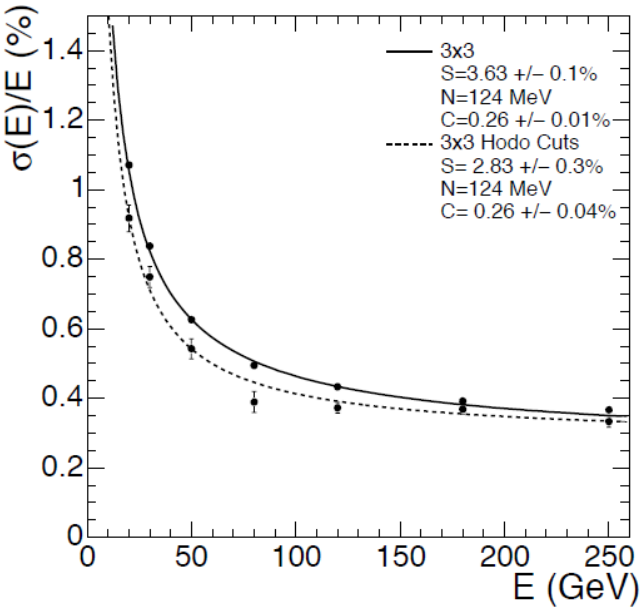

Figure 2.11: ECAL supermodule energy resolution. - ECAL energy resolution as a function of electron energy. The upper and lower points correspond to different data taking conditions. The parameters of the fit can be seen in the figure legend.

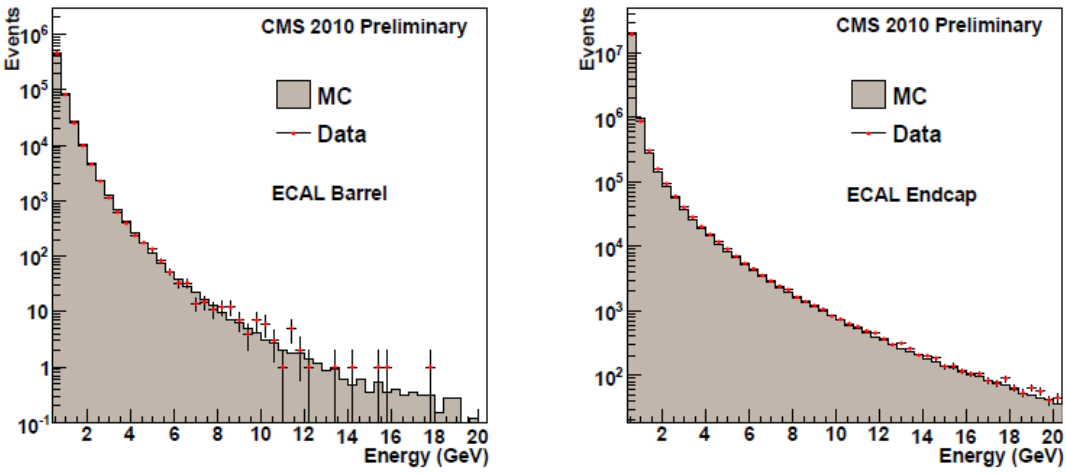

Figure 2.12: Energy spectra of ECAL. - Energy spectra of the individual channels in the barrel (left) and in the endcaps (right) from 7 TeV collision events (29).

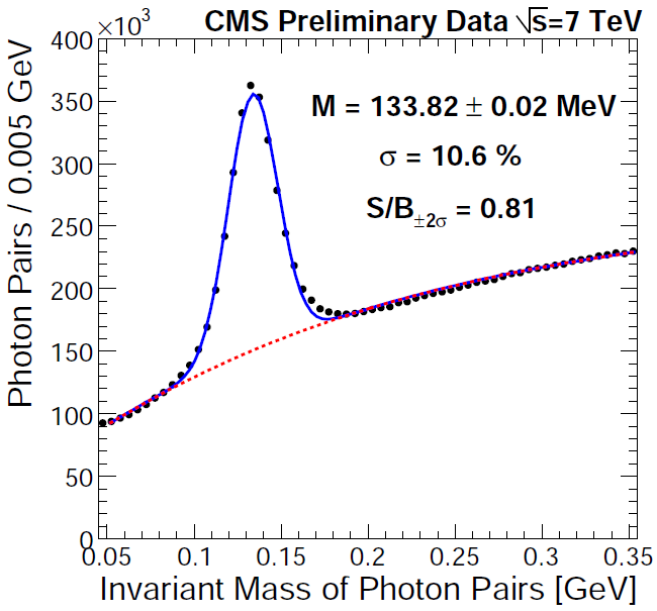

Figure 2.13: Invariant mass of π^0 candidates. - Mass spectrum of reconstructed pairs of photons showing a clear peak corresponding to the π^0 mass (29).

since it is located inside the solenoid, surrounding the ECAL. Hence, the HCAL design maximizes material inside the magnet coil in terms of radiation lengths. The material used as absorber is brass due to its short interaction length, being easy to machine and non-magnetic. The active medium uses plastic scintillator tiles connected with wavelength-shifting fibers (WLS) and hybrid photodiodes to detect the emitted light.

The HCAL consists of 4 modules: the inner barrel, the outer barrel, the endcaps and forward calorimeter.

The barrel calorimeter (HB) covers the pseudorapidity region $-1.4 < |\eta| < 1.4$. It consists of 2304 towers with a segmentation $\delta\eta \times \delta\phi = 0.087 \times 0.087$. It is constructed out of flat brass absorber plates, parallel to CMS z axis, interleaved with scintillation plastics.

The outer part (HO) covers the region $|\eta| < 1.26$. It contains scintillators with a thickness of 10 mm. The tiles are grouped in ϕ matching DT segmentation (angular sectors of 30 degrees in ϕ). The outer part increases the hadron calorimetry to 10 interaction lengths. MET resolution is also improved by HO. It is divided in 5

rings. The central one, in which the absorber length is minimal, uses 2 scintillator layers (18 cm thick). The other rings have 1 layer each.

The hadron endcap (HE) covers the pseudorapidity region $1.3 < |\eta| < 3.0$, a region in which about 34% of the particles from the primary vertex are produced. It consists of 2304 towers. They are distributed in 14 different η regions, covering 5 degrees in ϕ each the outermosts and 10 degrees the innermosts.

The hadron forward (HF) gives coverage in the $3 < |\eta| < 5$ region. Its design is determined by the high particle fluxes in this pseudorapidity region. It consists of 900 towers and 1800 channels in the 2 modules. They are grouped in 13 η regions, covering each tower 10 degrees in ϕ (except the highest η tower, covering 20 degrees). The front face is situated at 11.2 m from the interaction point. The absorber is 1.65 m deep. The Cherenkov light emitted is transmitted by quartz fibers to reach the photomultipliers.

The performance of the hadronic calorimeter was studied with simulation and with test beams before its installation underground (see Fig. 2.14). In Fig. 2.15, the p_T resolution as a function of the p_T and for three different jet reconstruction algorithms is shown (31) in simulation. In data, jet p_T resolution is shown in Fig. 2.16, for γ+jet and dijets.

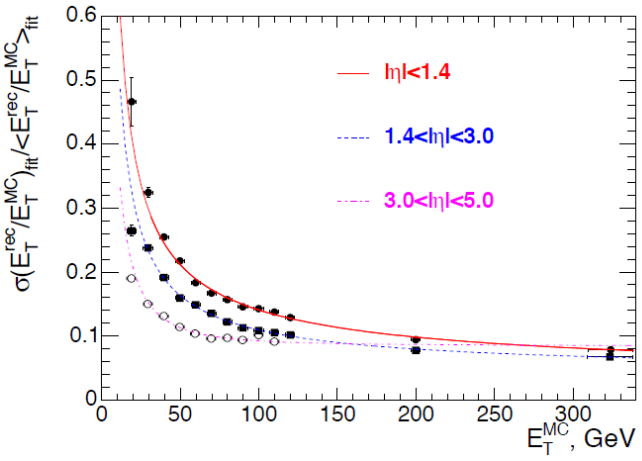

Figure 2.14: HCAL energy resolution - The jet transverse energy resolution as a function of the simulated jet transverse energy in 3 different η regions.

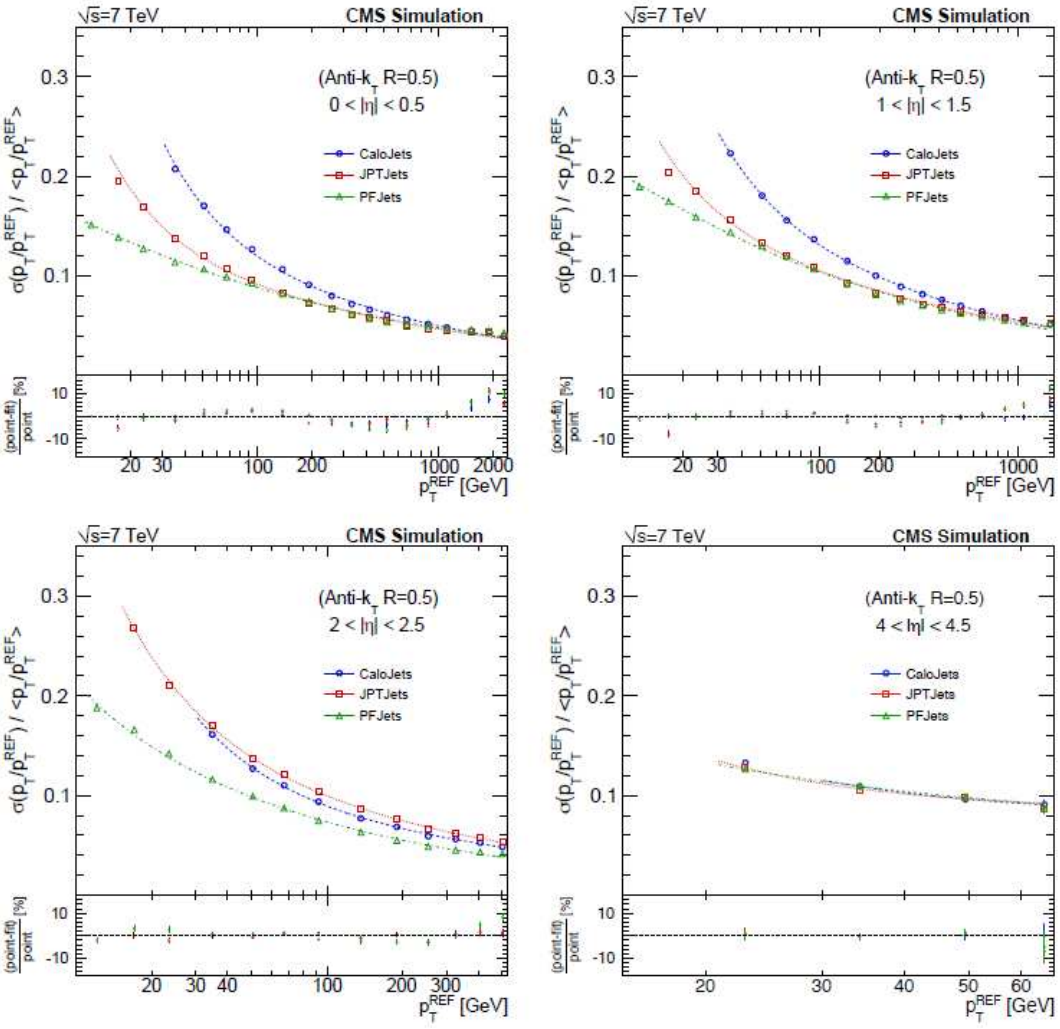

Figure 2.15: HCAL p_T resolution - p_T resolution as a function of the p_T is shown for four representative η ranges.

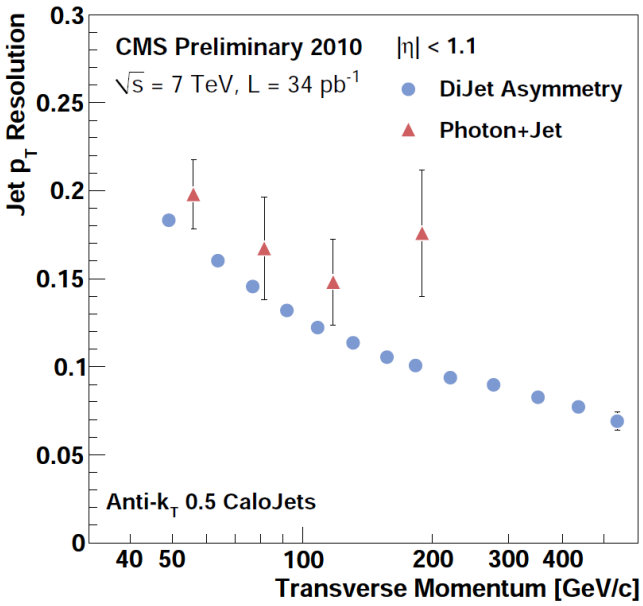

Figure 2.16: - Jet p_T resolutions from γ+jets (red triangles) and dijet asymmetry (blue points) measurements (32).

- **Muon system**. The aim of the muon system is the measurement of the trajectory of the muons and their identification, as well as to provide with seeds for the different levels of the trigger system. Three different detector technologies are used to detect muons: cathode strip chambers, resistive plate chambers and drift tubes. The importance of muons in CMS, a description of these three detectors, and calibration and parametrization results carried out are presented in the next chapter.

3

Muon detection and reconstruction in CMS

Muons are charged elementary particles belonging to the second lepton family. They are unstable particles, that decay in electrons and neutrinos with a mean lifetime of 2.2 μs (typical muons in high energy collisions travel more than 600 m before decaying) (5). The mass of the muon, 105.7 MeV, is 200 times bigger than that of the electrons. Muons could be thought of as heavy electrons. As a consequence, they do not emit as much bremsstranhlung radiation as electrons do. Thus, muons penetrate into matter further than electrons, being able to go through long distances before being absorbed.

3.1 Importance of muons in particle physics analysis

The importance of muons in the CMS experiment is patent in its name. The "M" of CMS refers to the "muon" particles, underlining the key role of this particle in the experiment. There are several reasons that make this particle so interesting from the physical analysis point of view. First of all, concerning detection, high energy muons are not absorbed by the detector as they are low interacting particles. This makes possible to study them in a low occupancy environment, since most other particles produced in the primary interaction are absorbed in the CMS calorimeters. As a consequence, physical processes involving muons in their final state are clean, and can be detected and triggered on with high efficiency (33).

Muon detection is non-destructive and thus by placing multiple layers of muon

detectors, an accurate and robust momentum measurement is achieved. The excellent transverse momentum resolution of muons in CMS and the almost 4π coverage of the detector guarantees a good reconstruction of any process related with muons.

Muons are interesting from the physics point of view. The clean muon detection provides experiments with a powerful tool, as many interesting processes contain muons in their final state. High transverse momentum muons come often from the decay of high mass particles. Since these particles are potentially interesting, muon detection is a reliable way to explore new resonances. Many of the interesting physical processes to be studied at the LHC can be studied in a muon channel:

- **Higgs boson.** As described in the previous chapter, the discovery (or rejection) of the Higgs boson is one of the fundamental goals of the LHC experiment (34).

  The Higgs boson, if it exists, will be detectable by its decay into less massive particles. Its decay into two Z bosons ($H \rightarrow ZZ^{(*)}$), in which both bosons decay into two muons is specially interesting. This final state with 4 high momentum, isolated muons, two positive and two negative, with a dimuon invariant mass near the Z peak, is one of the most promising channels due to the clean way the muons are detected (35) despite its low cross section compared to other possible decay channels.

  Other interesting channels in the Higgs boson search are $H \rightarrow WW \rightarrow l\nu l\nu$, $H \rightarrow WW \rightarrow l\nu jj$ and $H \rightarrow ZZ \rightarrow ll\nu\nu$. The first one is specially important in the [150 - 180] GeV Higss boson mass window, where the branching ratio for the golden channel (decay into Z bosons) is not significant enough. These four channels underline the importance of muon detection in almost the full Higgs boson mass spectrum.

- **High dimuon mass resonances.** Possible SM extensions predict the existence of new gauge fields. High dimuon mass resonances (Z', W') could appear as a consequence, confirming this prediction. The LHC will be able to study the existence of such resonances up to energies of the order of 4 TeV. This could be achieved by means of the measurement of high momentum muons, following decays like $Z' \rightarrow \mu^+\mu^-$ or $W' \rightarrow \mu\nu$.

- **Supersymmetry.** Supersymmetry (SUSY) is one of the most famous SM extensions. Muons are a key part in this search. Due to their properties, they are important in the discovery of the "sparticles" (squarks, gluinos, sleptons...) and the measurements of their properties. Examples are a possible decay of neutralinos in dimuon pairs, or the decay of the Minimal SuperSymmetric Model (MSSM) Higgs bosons, similar to that of the SM.

- **Electroweak physics.** Electroweak physics is important, especially in the early phases of the LHC operation. Their study contribute to improve the knowledge on the proton composition, are sensitive to new physical signatures, are background of many searches and are useful to test the SM in a new energy regime. Electroweak physics are mostly studied in the electronic and muonic channels, thus highlighting the importance of muon detection in this field.

- **Top physics.** Top quark are produced massively in LHC proton-proton collisions. They decay almost exclusively in a W boson and a b quark. The study of the muonic decay of the W is preferred over the hadronic one, due to the clean muon detection and the relatively easy QCD multijet background reduction. Muons are also important in the b-tagging methods, by analyzing the properties of the muon inside the jet.

3.2 CMS muon system

As briefly described in the previous chapter, CMS muon system is required to be accurate in the muon transverse momentum measurement, with excellent muon identification, over a wide region of momenta in the pseudorapidity region covered. Muons transversing CMS are first measured in the tracker as charged particles (innermost part of CMS) and then in the muon system (outermost part of CMS).

CMS uses three different technologies to detect muons (33).

The choice of each technology is mainly driven by two factors: the flux of particles and the magnetic field in the region considered. In the barrel region ($|\eta| < 1.1$), due to a low arrival flux of particles, a relatively slow detector can be used ($\approx$ 100 ns). The uniform magnetic field in this region allows the use of Drift Tubes (DTs). In the endcaps, where the flux of particles and the magnetic field is higher Cathode Strip

Chambers (CSC) are deployed, covering the region up to $|\eta| < 2.5$. A third technology, resistive plate chambers (RPC), is used in both barrel and endcaps for timing purposes. They are fast detectors with good time resolution and coarser position resolution than the other 2 kind of muon detectors. They are important devices to correctly identify the bunch crossing from which the detected particles come and for triggering purposes. The layout of the muon system in one quarter of the longitudinal plane of CMS can be seen in Fig. 3.1.

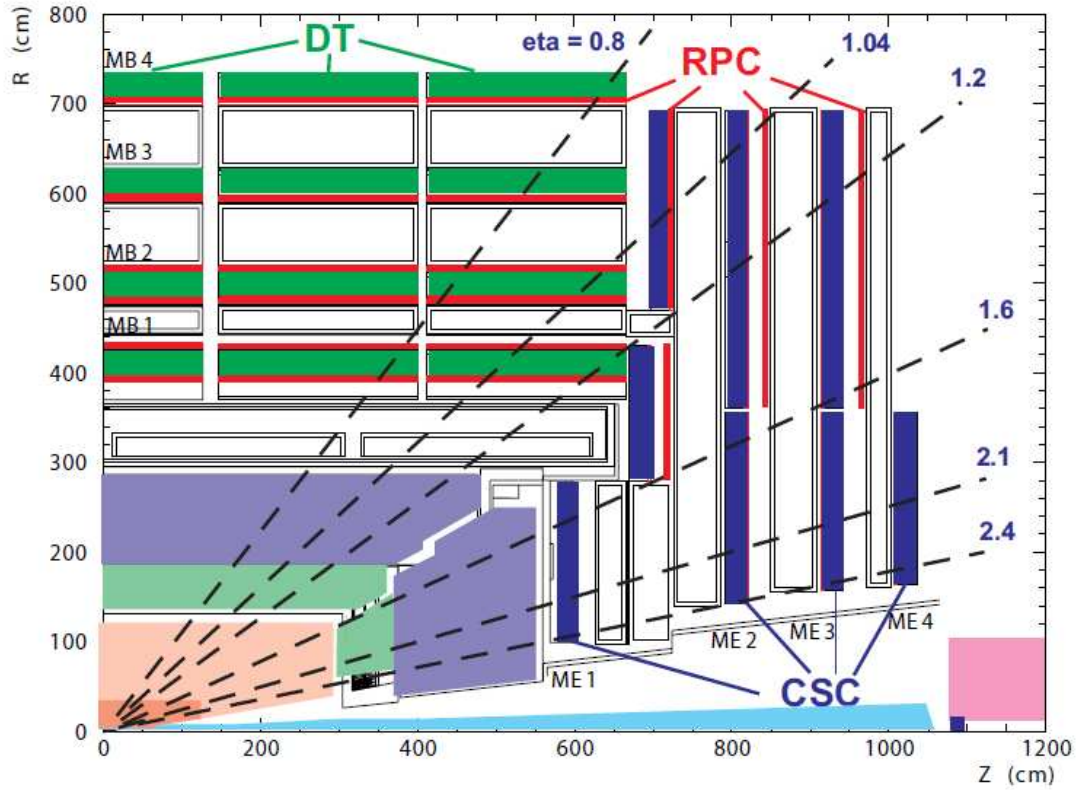

Figure 3.1: Layout of the muon system - Distribution of muon chambers in a quarter of the longitudinal plane of CMS. DT and RPC chambers are used in the barrel, whereas CSCs and RPC ones in the endcaps

The muon system is also a key part of the trigger system. It operates within the first level, providing two independent and complementary sources of information (CSC or DT, and RPC). The high number of chambers in the muon system, and the use of two different technologies to provide information make the muon trigger to be a robust, precise and reliable source of information.

All muon detectors in CMS are gaseous detectors. When a muon traverses the gas it may produce the ionization of nearby atoms. The difference of electrical potential applied in the chamber accelerates the free electrons creating an avalanche. The cloud of electrons reaches the anode wire, being detected as an electrical signal.

3.2.1 Cathode Strip Chambers.

These are the detectors chosen to detect muons in the forward part of CMS. CMS uses 468 CSC to form the Muon Endcap (ME) system. The CSC chambers are trapezoidal detectors. Each of them consists of 6 gas gaps, with a plane of radial strips and a plane of anode wires running perpendicularly to the strips (see Fig. 3.2). While electrons are accelerated towards the anode wire, ions go to the cathodes, measuring 2 coordinates in each plane. To avoid gaps in the muon system, most of the chambers are overlapped in the azimuthal angle. By adding information of the different planes of a chamber, a 3D reconstruction of the original trajectory is built. This information is also used for the Level-1 Trigger. The spatial resolution is about 200 μm and the angular resolution, 10 mrad.

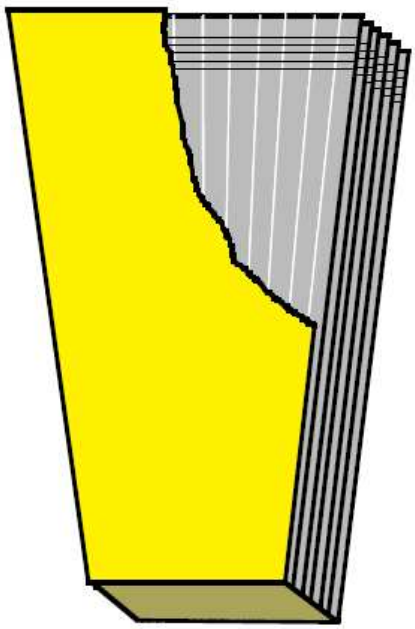

Figure 3.2: CSC representation - Schematic view of an endcap CSC: a six plane chamber of a trapezoidal shape with strings running radially and wires running across.

3.2.2 Resistive Plate Chambers.

They are used for timing purposes due to their fast response. They consist of 2 parallel planes of plastic material, with a 2 mm gap, with a high voltage applied between them (see Fig. 3.3).

RPCs can either operate in streamer or avalanche mode. In CMS this second mode is chosen due to the high flux of particles expected (10 kHz/cm^2).

In the barrel, there are 1 or 2 RPCs placed coupled to the Drift Tube chamber, depending on the station. In each of the two endcap rings 36 RPC chambers are used.

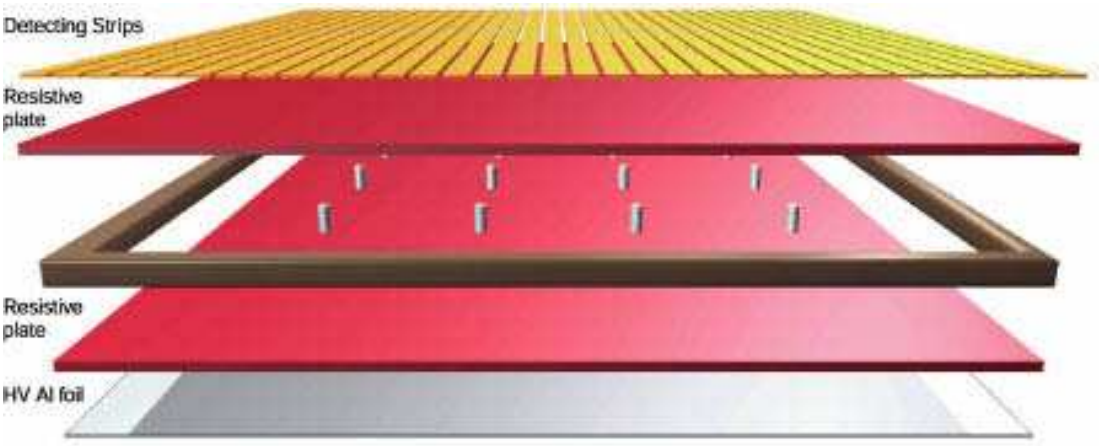

Figure 3.3: RPC representation - RPC detectors in CMS are made of 2 parallel planes of plastic filled with gas.

3.2.3 Drift Tube Chambers.

Since these detectors are a capital part of the thesis, they will be explained with more detail. The barrel muon system at CMS consists of 250 drift tube (DT) chambers allocated in 5 iron wheels, which serve to guide the return of the magnetic field lines of the solenoid, named YB-2 consecutively to YB+2. As a consequence, drift tubes are under the influence of the residual magnetic field created by the solenoid. Each wheel is structured in 12 sectors named counter clockwise from S1 to S12, each one covering 30 degrees in the $r-\phi$ plane (see Fig. 3.4). Four stations are placed in concentric rings around the beam line in each sector to measure the muon trajectory. They are labeled MB1 to MB4 from the innermost to the outermost ones.

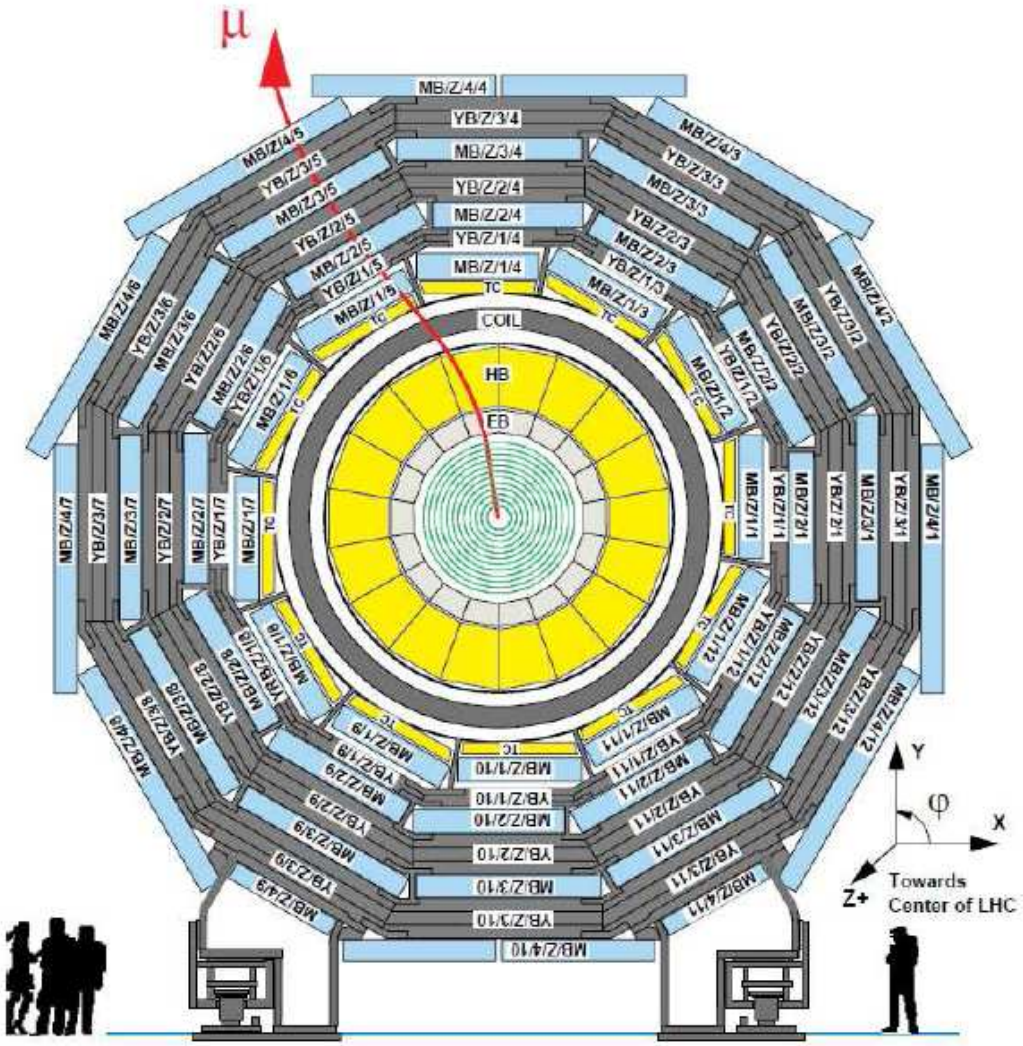

Figure 3.4: Schematic view of the CMS detector. - Transverse view of CMS. The barrel muon detector elements are denoted as MBZ/N/S where Z is the wheel number, S one of the 12 sectors and N the station number.

The basic detection unity is the drift cell, shown in Fig. 3.5. These drift tubes are put together in rows (called layers). 4 rows are glued one on top of each other to form a so called SuperLayer (hereafter SL), measuring either the $r-\phi$ coordinate ($SL-\phi$) or the $r-\theta$ coordinate ($SL-\theta$) depending on the wire orientation. These 4 rows are placed staggered to diminish the probability of losing a muon. Stations MB1, MB2 and MB3 combine 2 $SL-\phi$ and 1 $SL-\theta$, whereas MB4 measures only the $r-\phi$ plane (see Fig. 3.6). In order to confer rigidity to the ensemble a non-sensitive structure, called honey-comb, is inserted between the SuperLayers.

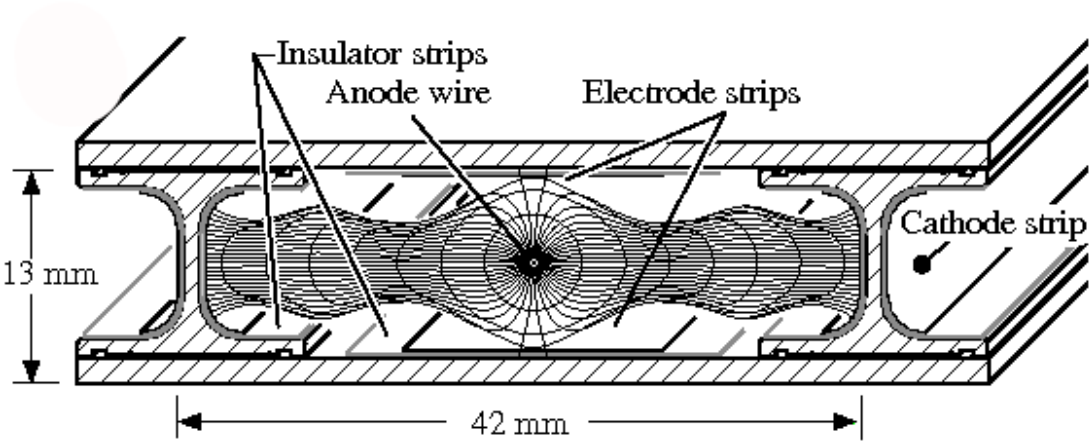

Figure 3.5: CMS barrel muon drift cell - Representation of CMS drift tube chamber detection unity.

The drift cell is a 2.5 meters long tube of rectangular section (42 mm long, 13 mm high). Right in the middle, an anode wire runs along the tube. At a distance of 21 mm, the cathode strip runs in the same direction. The drift cell is filled with a gas mixture (15% CO2, 85% Ar) optimized to detect a particle going through it. The 2.3 kV/cm electric field created inside the tube makes the free electrons ionized after the passage of the muon to drift towards the wire. The time it takes the electrons to reach the wire is measured as an absolute time (TDC counts). This time is converted into the drift time using the calibration procedure detailed in the following sections. This drift time measured is related to the relevant quantity, the distance, to calculate the passing point of the charged particle inside the cell. This is done using the electron drift velocity in the cell. The electric field inside the cell is required to be highly uniform in order to have a constant drift velocity, and avoid any non-linearities. In one of the two ends of

the drift tube, the high voltage is provided. This end will be called along the text as *high voltage* side (HV). At the other end, the read-out electronics are placed. It will be called hereafter the *front-end* side (FE).

Each drift tube gives information of one point in the muon trajectory. When a muon crosses the barrel region, the DT chambers provide up to a total of 44 points (32 in the $r-\phi$ plane and 12 in the $r-\theta$ plane) which are used to build the muon trajectory. This information is later combined with the data coming from the tracker to conform the complete muon trajectory. The way the local information given by a DT chamber is transformed in a full muon path in the barrel region is hierarchical, building more complex objects as we associate information coming from neighboring detectors. All the hits coming from a superlayer are fitted to a straight line producing a so called 2Dsegment (two-dimensional information). All the 2Dsegments in a chamber are then used to create a 4Dsegment, with 3-dimensional information. They are, in turn, again associated to obtain the highest order object using DT chambers alone: the Stand Alone Muon Track (STA) (36). To build the STA track it is necessary to bear in mind the effect of the magnetic field on the muon path. Since the magnetic field inside the drift cell is negligible, the path within a muon chamber can be considered as linear, and reconstructed fitting a straight line. However, in the iron that separates the drift chambers in a same sector, the magnetic field is of the order of 2 T, bending the muon trajectory. A dedicated algorithm reconstructs the complete muon trajectory considering the muon scattering in this dense material and the curvature of the muon due to the magnetic field.

The muon chambers were tested and commissioned in the laboratory using cosmic rays. Then, they were assembled at CERN and tested both in the surface and in the cavern. To check the synchronization with the other detectors, the chambers were tested using cosmic muons with the full CMS in operation mode. These tests were performed with the magnetic field switched off and at 3.8 T. The studies presented in this Chapter correspond to the calibration of the detector carried out using all CMS subdetector (both at 0 and 3.8 T). To do so, cosmic muons were detected during several data taking campaigns. CRUZET (Cosmic Run at Zero Tesla) gives name to a series of campaigns in which the magnetic field of the solenoid was switched off. CRAFT08 (Cosmic Run at Four Tesla 2008) is another of these multiple tests that the CMS subdetectors have undergone. It was the first global run with the magnetic field risen

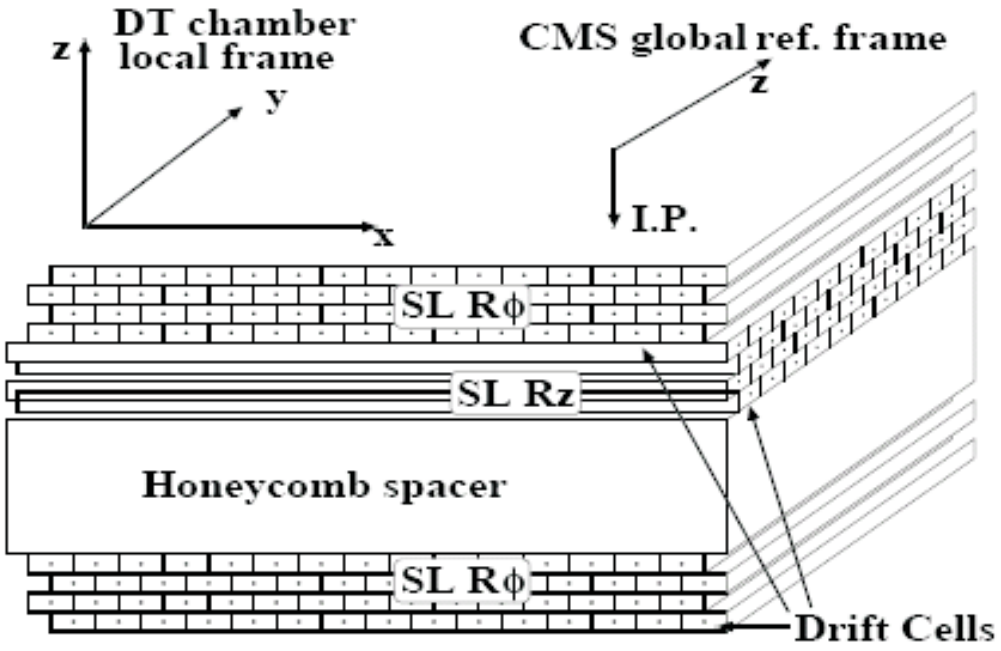

Figure 3.6: CMS DT chamber - A DT chamber consist of 2 ϕ SL, a non-active structure (honey-comb) to give rigidity to the ensemble and, for MB1, MB2 and MB3, a θ SL.

at its design value. During this data taking period, spanning around one month of continuous running, more than 300 million cosmic muon events were registered. The details of some of these runs are presented in Tab. 3.1. The conditions of the detector (magnetic field) during these data taking periods is shown in Fig. 3.7.

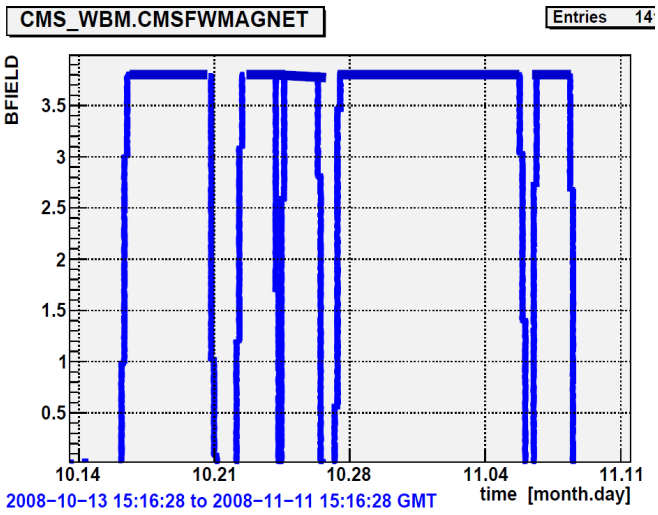

Figure 3.7: Magnetic field - Status of the solenoidal magnetic field in CMS during the period in which the CRAFT08 was taken. Data with and without magnetic field were taken for CRAFT08.

These tests served to characterize the response of the drift tubes. Several studies were carried out using cosmic muons: the local reconstruction in the drift chambers of the muon tracks, the angular and spatial resolution in the different steps of the reconstruction, charge misidentification rate, DT trigger performance, reconstruction efficiencies and calibration and drift velocity measurements are the most important (37).

The spatial and angular hit resolution depend on the drift velocity in the detector and a good knowledge of it is needed. Consequently, a deep understanding of the different effects that may have an impact on the drift velocity determination is essential, in order to keep the resolution as high as possible. Drift velocity in the DTs is studied in section 3.

CMS muon system is required to be a robust, accurate system. In the barrel region, DTs should detect and reconstruct the incoming muon with high efficiency. The reconstruction efficiency of this detector is studied in section 4.

# RUN	B Field	# events analyzed	Data Taking
57553	0 T	508 705	CRUZET
57539	0 T	436 392	CRUZET
557498	0 T	154 364	CRUZET
67818	3.8 T	20 303 938	CRAFT
67838	3.8 T	34 316 877	CRAFT
68021	3.8 T	35 822 927	CRAFT
68141	3.8 T	3 520 000	CRAFT
70147	0 T	960 000	CRAFT
70170	0 T	1 300 000	CRAFT
70675	0 T	2 470 000	CRAFT

Table 3.1: CRAFT08 runs analyzed in this study.

3.3 Determination of the drift velocity

The method used in this study to measure the drift velocity is called *meantime method*, and it is based on computing the maximum drift time within a cell.

For every 3 consecutive layers in a SL (for example layers i, j and k) the value

$$T_{mean} = (T_i + T_k)/2 + T_j \tag{3.1}$$

corresponds to the time the ionization cloud takes to reach the wire when the muon crosses the cell next to the cathode (see Fig. 3.8) and it is called T_{mean}. In this formula, T_i is the drift time on a cell in the i layer. The meantime is the maximum drift time in a cell, for muon incident angles below 30 degrees, given the geometry of the drift cell and their displacement inside a SL, shown in Fig. 3.8. This value is of the order of 400 ns.

Provided the electrical field is uniform enough, the drift velocity can be taken as constant inside a cell and can be easily computed as half the cell size over the T_{mean},

$$v_d = \frac{21000}{T_{mean}} (\mu\text{m/ns}) \tag{3.2}$$

Time given by the read-out does not only account for the drift time but also for other contributions as the propagation delay along the cables and the trigger latency (the delay produced by the trigger system in the data acquisition). The conversion of

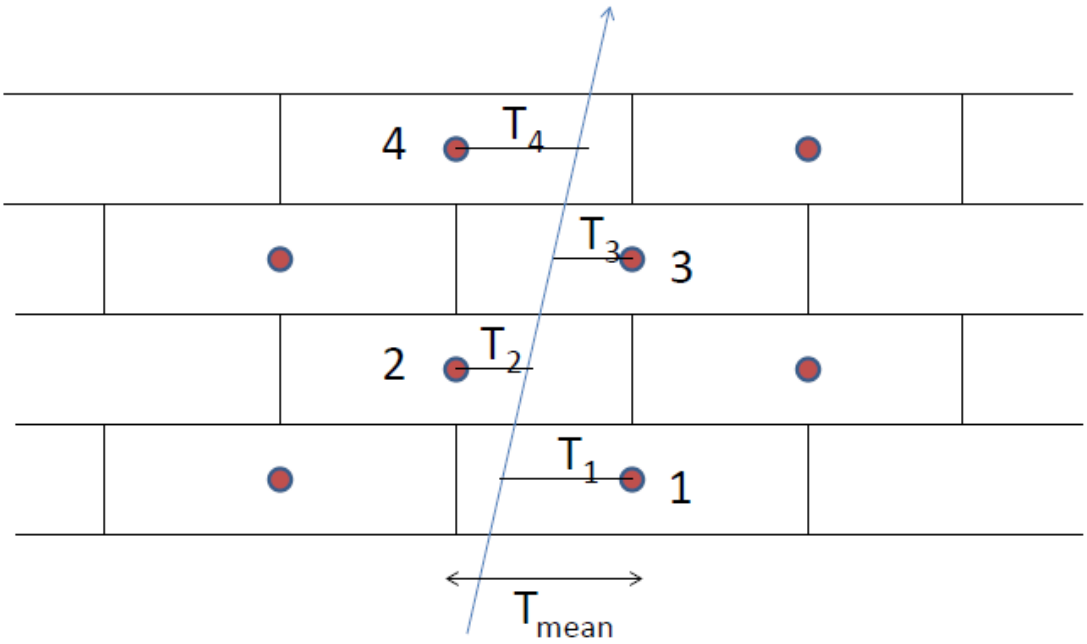

Figure 3.8: Meantime method computation in a SL - This method is based in the determination of the maximum drift time on a cell.

this raw time (TDC counts) into a drift time is done through a dedicated calibration process.

3.3.1 Calibration of the drift tube

The *meantime method* relies on an accurate calibration of the SL, which essentially gives the time origin, signaling the passage of a muon through the detector. There are several sources that contribute to this raw time. Small differences among the propagation time in different cells in layers are taken into account in the so called relative t_0 (few ns). This time is computed sending LED pulses to every cell, in dedicated calibration runs. In practice, in the DT chambers, this procedure is applied for a group of cells (2 layers of cells or a SuperLayer). Once the relative t_0 values are taken into account, a raw time histogram or time distribution (*tbox*) can be plotted for a group of cells (Fig. 3.9 (left)). The time offset, in this case around 1400 ns, is known as global t_0 or *time pedestal* and it is related to the trigger latency and the cabling from the detector to the readout units. Both relative and global t_0 factors integrate the final t_0 (37). A typical time box after calibration can be seen in Fig. 3.9 (right). The time pedestal or t_0 corresponds,

thus, with the time read when a muon passes close to the wire. This value is computed for a reference run and stored for each cell in a data base. After calibration procedure, the intrinsic drift tube resolution is of 4-5 ns.

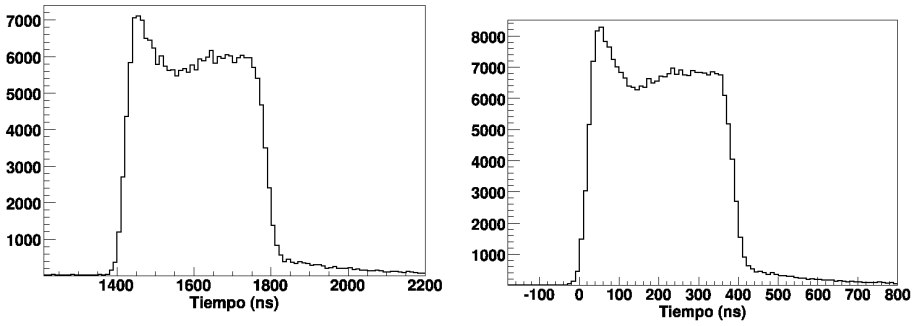

Figure 3.9: Time box - A typical time distribution (or tbox) (left) before calibration, (right) after calibration.

In collision mode, muons are expected to arrive every 25 ns. Windows of 25 ns width are then used to distinguish among different bunch crossings. However, when collecting cosmics, the muon arrival time to the drift chambers is not known a priori and thus, is not synchronized with LHC clock. This fact adds an uncertainty of 12.5 ns in the muon arrival time that propagates to the drift time determination and therefore affect the drift velocity computation and the chamber resolution.

In this section, 4 different runs, belonging to the CRAFT08 final days, are chosen to analyze the drift velocity with and without magnetic field. The details of these runs are presented in Tab. 3.1. Some runs from CRUZET4 (previous to CRAFT08) were also used for specific studies. All the results presented correspond to one of the the SL-ϕ as representative.

The value of T_{mean} has been computed for each $SL-\phi 1$ of MB1, MB2 and MB3 stations in sectors 4 and 10 of each wheel and in every run, by means of equation 3.1. Sectors 4 and 10 are chosen since they are the horizontal ones. Due to the fact that cosmic rays arrive mostly in the vertical direction, the horizontal chambers will reconstruct a bigger amount of cosmic muons. An example of a T_{mean} distribution is plotted in Fig. 3.10. The graphic on the left shows the T_{mean} distribution obtained after the t_0 calibration. To compute the right plot, a correction on the arrival time is applied to make the plot look like the one we have in proton-proton collisions. In this case, the

arrival time is introduced as a new parameter in the fit to the muon hits, improving the resolution. In both cases, the distribution is fitted to a gaussian function, and the mean value is used to compute the drift velocity by means of equation 3.2. In Fig. 3.11 the drift velocity for all $SL-\phi 1$ in every MB1 station of every sector in CMS is plotted, without the arrival time correction (left), and with it (right). The plots including the arrival time correction are shown as an illustration of the collision mode plots and this method is not applied on the rest of the Chapter. In order to quantify the systematic errors when applying this method the T_{mean} using layers 1, 2 and 3 of a SL and layers 2, 3, and 4 of the same SL is computed. The r.m.s. of the difference distribution between these two quantities gives an idea of the error associated with this method. An error of 0.12 μm/ns and 0.04 μm/ns is obtained in both cases (without and with correction respectively). However, this procedure used to obtain systematic errors is not sensitive to a bad calibration, as it affects equally T^{123}_{mean} and T^{234}_{mean}.

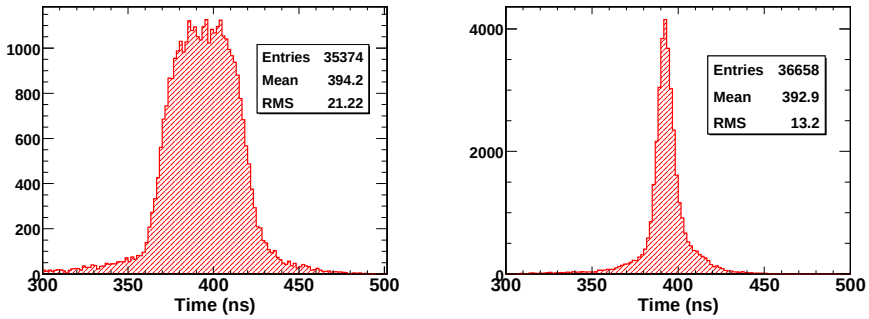

Figure 3.10: Tmean distribution - T_{mean} distribution for the W-2 S4 MB3 station (run 70147) (a) without and (b) with t_0 *event-by-event correction*.

It is important to notice that the drift velocity absolute magnitude is very sensitive to the t_0 determination procedure. This is the reason why the attention of this study is mainly focused in the drift velocity variation in different cases and not in its actual value.

In the following subsections the different factors that may affect the measurement of the drift velocity are revised.

- The muon incidence angle on the chamber is a factor to bear in mind. CMS is designed to optimize detection of muons with a maximum angle of 30-35 deg.

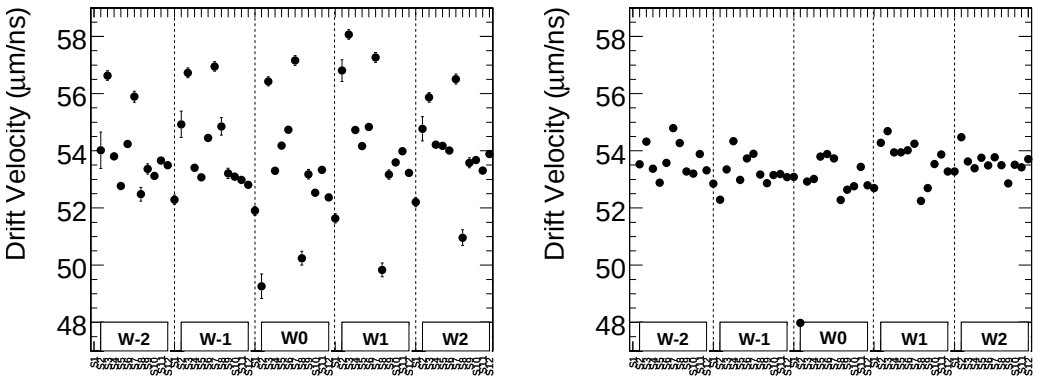

Figure 3.11: Drift velocity - Drift velocity for every $SL-\phi 1$ in MB1 stations in a run without magnetic field (a) without and (b) with t_0 *event-by-event correction*.

These are the conditions in the nominal p-p collision mode. Muons with higher incidence angle will create a different pattern of hit cells by ionized electrons affecting the measured distance.

- The length of the drift tube, 2.5 meters long, will introduce additional indeterminations. The time taken by the signal created after the avalanche to propagate along the wire will be added to the drift time, making measured distances longer. The total drift time will differ in a few ns, depending on where it was produced along the wire.

- The drift chambers are placed inside the CMS magnet return yoke, thus under the influence of the residual magnetic field created by the solenoid. This field varies considerably depending on the chamber position, as can be seen in Fig. 3.12, its radial component reaching around 0.8 T in the innermost stations (MB1) near the endcaps. As it will be further explained, this magnetic field affects the electron drift path.

3.3.2 Effect of the muon incidence angle in the drift velocity computation

When a muon traverses a DT chamber with a certain incident angle, with respect to the normal incidence, the electron cloud created by ionization will be lead by electrons

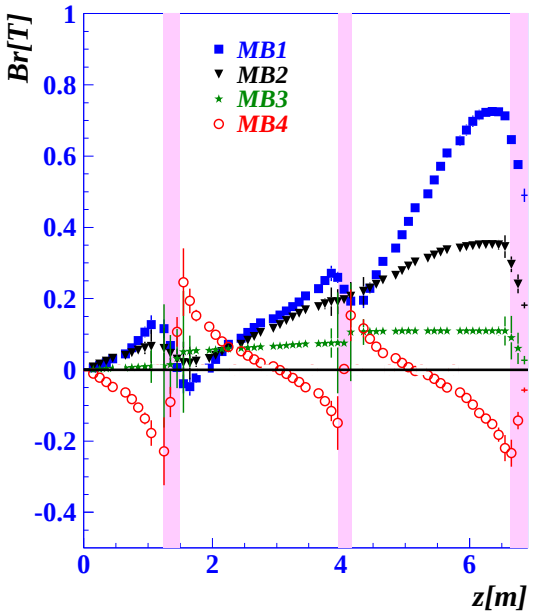

Figure 3.12: Radial magnetic field component along the CMS global z coordinate. - In the external parts CMS for MB1 stations a non negligible residual magnetic field is observed.

closer to the wire. In this case, the drift time computed will be smaller than in the case of vertical incidence, and thus the drift velocity will appear to have bigger values.

This effect, which is of the order of 3% for high incidence angle (greater than 20 degrees), is clearly seen for a 0T run. For that purpose, a group of CRUZET4 runs (57553, 57539, 57498) has been studied, dividing the angle spectrum in 4 intervals: muons with incidence angle between 0 and 5 degrees, 5 and 15, 15 and 25 and finally higher than 25 degrees.

An apparent drift velocity increase when increasing the incidence angle is observed in Fig. 3.13 for the $SL - \phi$. The drift velocity is plotted using data from CRUZET4, once CMS was installed inside the cavern. A good agreement between this plot and previous studies with the detector in surface is observed [**REF**]. The drift velocity variations reach around 3% for high incident angles.

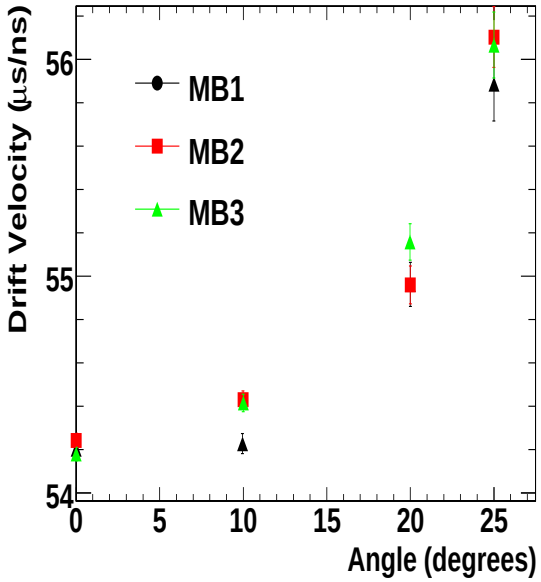

Figure 3.13: Drift velocity versus the incidence angle -. Drift velocity versus the muon incidence angle with data from CRUZET4 runs.

3.3.3 Effect of the signal propagation delay along the DT wire

The time stored by the DT electronics after a muon crossing contains not only the drift time and trigger latency but also the time the signal takes to reach the front-end.

Whereas trigger latency effect is removed when carrying out the calibration process, this latter contribution remains in the final time used in the reconstruction. In spite of the high speed of the signal propagation along the wire, the 2.5 meters long wire is enough to make this effect noticeable when comparing a muon crossing the DT near the front-end and other from the other side. The delays for two muons separated 1 meter along the wire direction is of the order of 3 ns. It is sizable compared with the intrinsic time resolution in the drift cell (4-5 ns) and consequently observable.

In order to study this effect, the drift tubes are divided in 5 regions in the wire direction, as can be seen in Fig. 3.14. For each single region, the T_{mean} is computed and the drift velocity is extracted. To know the region the muon traversed we make use of the information coming from the $SL - \theta$.

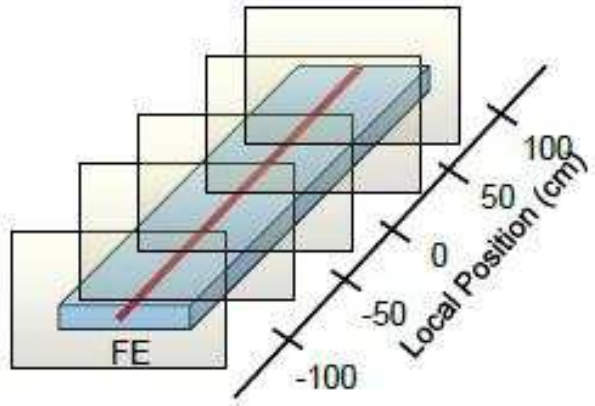

Figure 3.14: Chamber division - Graphical description of the chamber split carried out to study the effect of the signal propagation in the wire and the effect of the magnetic field in the drift velocity within a chamber.

In order to study this effect independently from the magnetic field effect, several runs taken at 0 T from the final part of CRAFT are chosen. Details from the runs analyzed in this section are shown in Tab. 3.1.

The drift velocity computed for the stations W-2 S4 MB1, MB2 and MB3 is presented in Fig. 3.15 for each of the 5 regions defined. The read-out electronics are placed in the negative extreme of the local X coordinate, where the computed drift velocity is maximal, since the length to be traveled by the signal is minimal. The further a muon crosses the cell towards the opposite side, the bigger the distance to be traversed by the signal, decreasing the computed drift velocity.

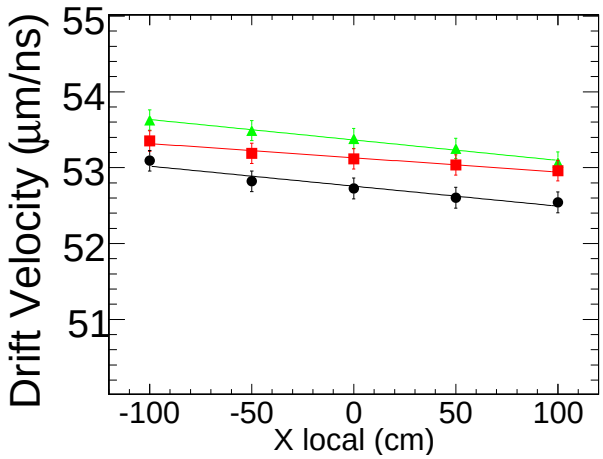

Figure 3.15: Drift velocity versus chamber local X coordinate - Drift velocity computed for stations W-2 S4 MB1 (black dots), MB2 (red squares) and MB3 (green triangles) versus the chamber local X coordinate. Similar plots were obtained for the rest of the CMS DT chambers with analogous results.

The velocity of the signal propagation along the wire can thus be computed as it is directly related to the slope of the linear relation in Fig. 3.15, through the expression

$$v_w = \frac{2v_d}{-T_{mean} \cdot m(v_d)} \tag{3.3}$$

where T_{mean} is the mean value of the T_{mean} distribution in each region, v_w is the signal propagation velocity, v_d the average drift velocity in the chamber, and $m(v_d)$ the fitted value for the slope in the drift velocity versus X coordinate plot. The average value for the signal propagation velocity is compatible with the speed of light.

3.3.4 Effect of the magnetic field in a ϕ-Superlayer

The magnetic field is expected to affect the drift path of the electrons in their way to the anode. The CMS superconducting solenoid provides a magnetic field of 3.8 T, essentially parallel to the Z-axis in the global CMS system of reference, inside the solenoid, with small radial and axial components in the DT chambers. This residual radial field, orthogonal to the electron drift velocity in a ϕ-superlayer, will give rise to a non-zero v_z component of the electron drift velocity, making the drift path longer. As a result, the apparent drift velocity (e.g. the one measured with the *meantime method*) will be reduced, in such a way that the higher the radial component of the magnetic

field, the lower the apparent drift velocity in that chamber. The axial magnetic field component is not expected to change notably the drift path, as it will be roughly 0 inside the DT chambers, giving rise to a 0 net force (only in a $SL - \phi$). The typical effect of the magnetic field on a CMS DT cell over the drift lines is shown in Fig. 3.16.

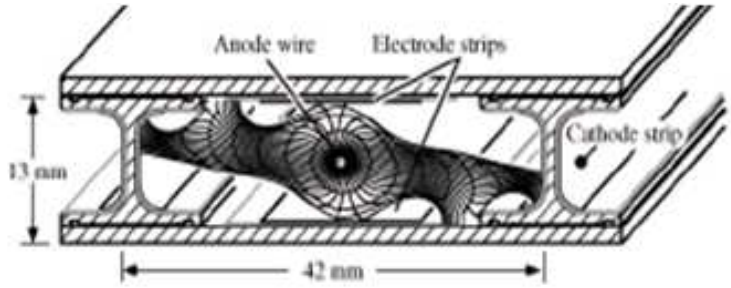

Figure 3.16: Effect of the magnetic field on the CMS cell drift lines. - The radial component in a $r - \phi$ SL makes the measured drift velocity to be smaller.

Fig. 3.12 shows the radial component of the magnetic field as a function of the distance to the center of CMS along the beam axis. The radial component is not uniform and ranges from 0 T to 0.8 T depending on the station and the wheel (z global coordinate) considered. As a result, the drift velocity may be different when comparing 0 T and 3.8 T runs, the effect becoming more evident in the stations where the radial component of the magnetic field is highest.

3.3.4.1 Effect of the magnetic field in the drift velocity computation for all CMS wheels

A comparison between drift velocities obtained in CRAFT runs (see Tab. 3.1) with and without magnetic field are presented in this section. In the B off case, data from the three runs analyzed, which share the same conditions, are accumulated in order to increase the statistics used. The drift velocities are computed for the superlayers $\phi 1$ in stations MB1, MB2 and MB3 from sector 4, in all the 5 wheels, using the *meantime method* previously explained. Hits from layers 1,2 and 3 (T^{123}_{tmean}) are used. The results are shown in Fig. 3.17. The errors in these plots are statistical and systematical, the latter computed, as previously explained, using the relative variation between T^{123}_{mean} and T^{234}_{mean}.

On the left plot of Fig. 3.17 the drift velocity for each station under study in both (B on/off) cases in sector 4 is presented. Comparing B on/off scenarios, a significant

decrease of the drift velocity when the magnetic field is switched on, is observed for both MB1 chambers from the outer wheels (wheels -2 and 2).

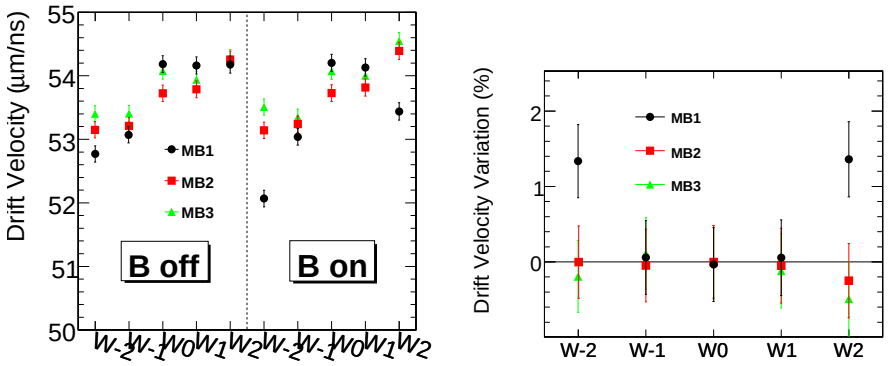

Figure 3.17: Drift velocity in a 0 and 3.8 T run -(Left) Drift velocity for superlayers $\phi 1$ in stations MB1 (circles), MB2 (squares) and MB3 (triangles) in sector 4 for each wheel. This plot is divided in two parts by a dashed line. Values on the left correspond to the 0 T case whilst right values are from the 3.8 T case. (Right) The variation between both magnetic field cases is shown: MB1 (circles), MB2 (squares) and MB3 (triangles).

In Fig. 3.17 (right) the percentile variation of drift velocity (B off case - B on case / B off case) is plotted, for the results presented in the upper plot.

The same procedure is applied for the rest of the CMS sectors, obtaining similar results in the overall effect of the magnetic field on the apparent drift velocity.

3.3.5 Effect of the magnetic field on the drift velocity computation within a chamber

As already shown in Fig, 3.12, the radial component of the magnetic field changes considerably along the CMS global z coordinate, mainly in the first station (MB1) of the outer wheels.

The values for the drift velocities presented in Fig. 3.17 for these chambers can be understood as an average value within the chamber.

In order to study the magnetic field effect with more detail, the chambers have been divided in 5 regions or *slices* along z (see Fig. 3.14) and the drift velocity computed in each one of them. This study has been done with the hits collected in the superlayer $\phi 1$ in stations MB1, MB2 and MB3 of sector 4, in Wheel -2. In the ϕ superlayers the

wires are disposed parallel to the z axis, with the front-end side placed in the outer part of each wheel. The effect of signal propagation (bigger apparent drift velocity for muons crossing near the front end side) acts in an opposite way with respect to the magnetic field effect. Results are shown in Fig. 3.18. where the drift velocity for each station of W-2 S4, split into 5 regions, is shown in 0 T (left) and 3.8 T (right) cases. In the 0 T scenario (absence of magnetic field) there is a decrease of the drift velocity when increasing the local X coordinate in all three stations (in other words, when moving from the front-end side of the chamber to the high voltage side), due to the time invested by the electronic signal to propagate along the wire.

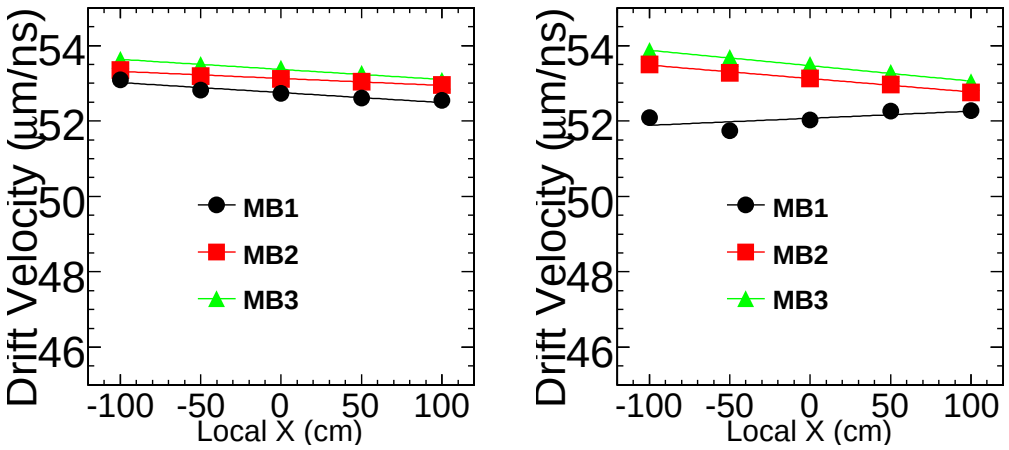

Figure 3.18: Drift velocity within a chamber - Drift velocity in $SL-\phi 1$ in stations MB1 (circles), MB2 (squares) and MB3 (triangles) in sector 4 W-2, in (left) a 0 T scenario, (right) a 3.8 T scenario.

In the presence of magnetic field (3.8 T scenario) the same description applies to drift velocities in MB2 and MB3, where the radial component of the magnetic field is almost negligible. In the case of the MB1 station, nevertheless, the effect of the magnetic field acts in the opposite way with respect to the signal propagation along the wire, and thus, the points are aligned in a 0 slope straight line, as a result of the counteract effect of both the magnetic field (decreasing the drift velocity near the front-end) and the signal propagation delay (increasing the drift velocity near the front end).

In Fig. 3.19 the same plots are shown, now for W-1. In this case no significant

differences can be appreciated between 0 T and 3.8 T scenarios. The radial component of the magnetic field in this wheel is not high enough to affect the drift velocity measurement, making 0 T and 3.8 T cases quite similar. The same behavior is observed in wheels 0 and 1, whereas results from W2 led to the very same conclusions extracted when studying Fig. 3.18.

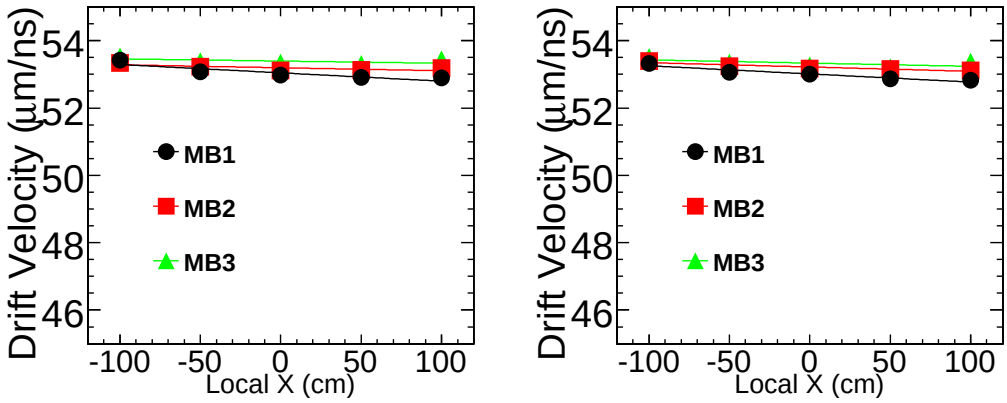

Figure 3.19: Drift velocity within a chamber - Drift velocity within the W-1 S4 $SL-\phi 1$ MB1 (circles), MB2 (squares) and MB3 (triangles) in (left) a 0 T scenario, (right) a 3.8 T run.

To summarize, in Fig. 3.20 the slope extracted from the previous plots (after a straight line fitting) is depicted for each wheel and station with and without magnetic field. As it has already been pointed out, only slopes from MB1 W-2 and W2 vary in both cases (B on/off). For any other station the slope is roughly constant and falls in a range between -0.008 and -0.006 μm/ns/cm as a consequence of the signal propagation along the wire.

To isolate the magnetic field effect we have compared the drift velocity in every z region with and without magnetic field.

Starting from the computation carried out in this section (see Fig. 3.18), we have worked out the variation for each region between 0 T and 3.8 T scenarios in all CMS wheels. The analysis is now extended to sectors S3, S4 and S5 and the results are averaged. In Fig. 3.21 this variation, (B on case - B off case) / Bon case, is shown for stations MB1, MB2 and MB3 for all the CMS wheels. The variation for chambers MB1, MB2 and MB3 from W0, W1 and W-1, and for MB2 and MB3 from W-2 and

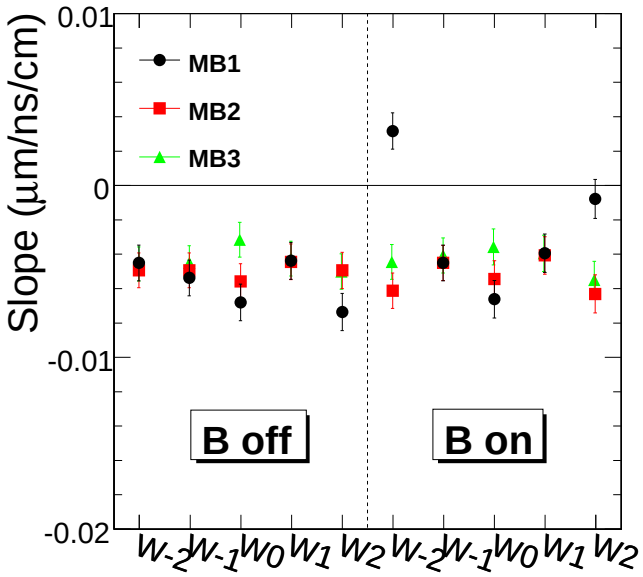

Figure 3.20: Signal propagation velocity along the wire - Slope from a linear fit of the drift velocity versus local z position in 0 T and 3.8 T scenarios for sector 5 of each wheel and stations MB1 (circles), MB2 (squares) and MB3 (triangles).

W2 are compatible with 0, as we expected since the radial component of the magnetic field is very low in these chambers. In MB1 from W-2 and W2, the variation behaves in the same way that the radial component of the magnetic field does, being maximal in the wheel outer region and minimal in the inner one (see Fig. 3.12).

These results can be compared with those drawn out during the MTCC [REF] observing a good agreement with those previous results.

3.4 Determination of DT reconstruction efficiency.

The reconstruction of the muon trajectory is done in several steps. Each of the steps is characterized by the efficiency of the reconstruction in that step. The hit reconstruction efficiency is bigger than 99.5% inside the cell and about 97.5% taking into account border effects of the drift cell. The 2D and 4DSegment reconstruction efficiencies of the DT chambers is reported in this Section, using CRAFT08 data (cosmic muons).

The computed efficiency is quantified as the fraction of muons crossing a DT chamber in which a 4Dsegment is successfully built. To determine whether a muon actually traverses a chamber, an extrapolation is performed making use of the hits recorded in the other stations. In order to guarantee that the extrapolation has been robust, some quality cuts are applied. The procedure is first carried out in top CMS sectors and then in the bottom ones. An alternative procedure to compute the segment efficiency is described in (37). Additional results obtained with CRAFT08 data, as the efficiency dependence with muon transverse momentum, are also reported here.

3.4.1 Data and sample selection

This study is carried out using data in conditions with and without magnetic field. Data from 5 different CRAFT08 runs (see Tab. 3.1) have been used: run 70147, belonging to CRAFT08 final days where the magnetic field was switched off; and runs 67818, 67838, 68021 and 68141 belonging to the last CRAFT08 period with the magnetic field on. The run at 0T is chosen to compare. The four 3.8T runs are shown to be good runs (good reconstruction in the whole CMS), and they had all the DT chambers operative (chambers from S7 and S1 are excluded, since they are vertical sectors, the amount of muons crossing this chambers and fulfilling the quality conditions are insufficient).

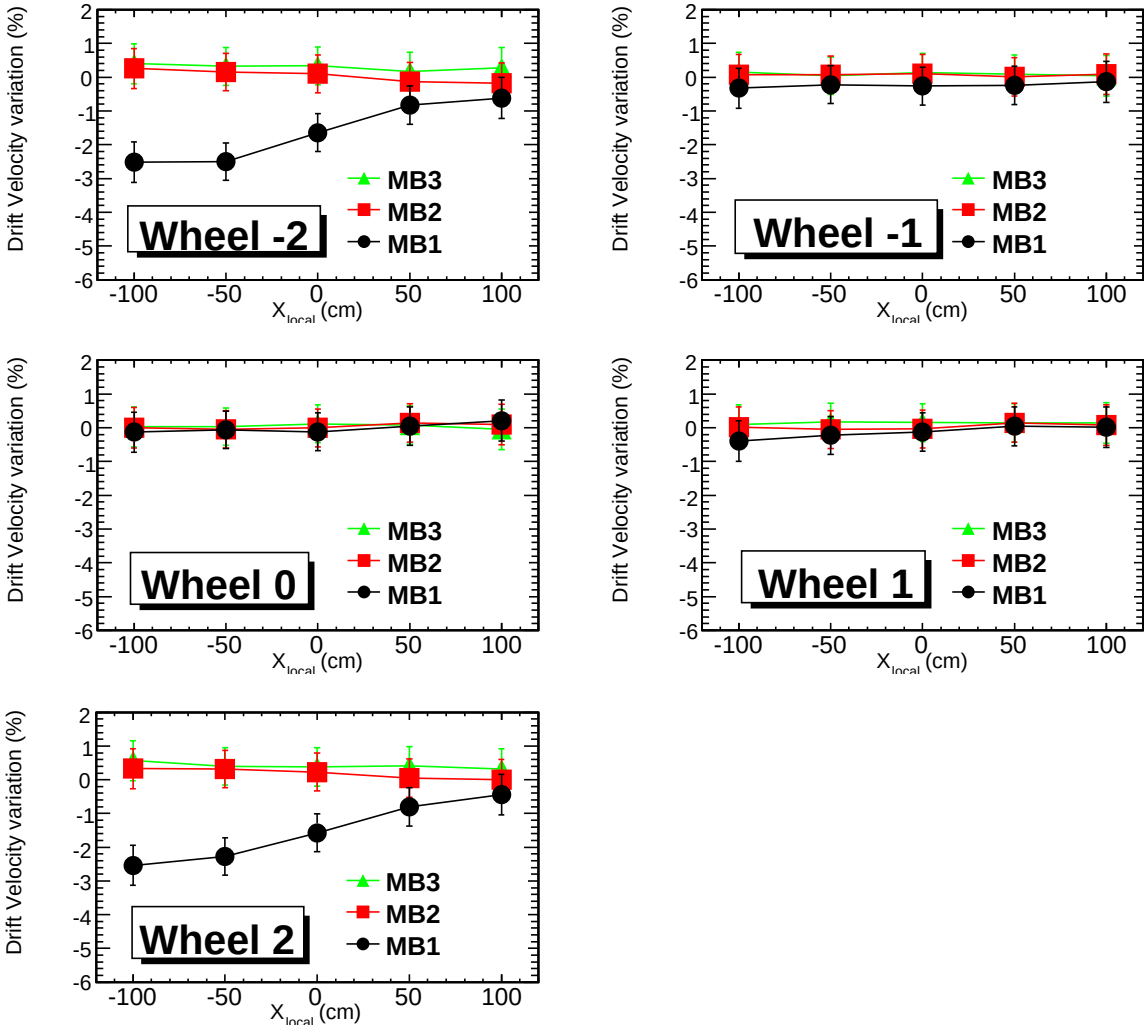

Figure 3.21: Drift velocity variation (CRAFT) - Drift velocity variation between the 0 T and the 3.8 T cases within the $SL-\phi 1$ of MB1 (circles), MB2 (squares) and MB3 (triangles) averaged over sectors 3, 4 and 5.

A total of 5 million events from each run were analyzed, which are, after the sample selection and quality cuts, enough to quantify the chamber efficiency.

The aim of this study is to measure the proportion of events where a particular DT chamber reconstructs a 4Dsegment when a cosmic muon traverses it. First events with a muon crossing each particular DT chamber are used. To determine the passing point of the muon we rely on the information delivered by the neighboring DT chambers. Hits from the other DT chambers in the same sector and wheel are used to build a muon trajectory. This track is then extrapolated to the test chamber to determine whether the muon crosses it.

To ensure a good extrapolation point we require the other 3 chambers (e.g. of MB1, MB2 and MB4 when testing MB3) to have a valid reconstructed 4Dsegment. Also, in order to simplify the 4Dsegment search and without generality loss, we will avoid using muons that cross different sectors and wheels, considering only muons that traverse the CMS bottom/top part within the same wheel and sector. Only events with one 4Dsegment per chamber are kept. In order to uncouple the 4Dsegment efficiency from any possible trigger effect, only events triggered by the opposite hemisphere to that of the test chamber are considered (events triggered by the top part to evaluate the bottom chamber efficiency and vice versa). We consider only events with a certain trigger quality: in the transverse plane HH or HL trigger quality (correlated) is required while in the θ plane any quality is accepted.

This selection process reduces the sample to around 10% of the initial one. Roughly, from trigger considerations one third of the initial sample is selected when studying the top sectors. And then, around another third of it fulfils the 4Dsegment requirement, having 3 4Dsegments in the same wheel and sector. Given the high statistics available, the selected sample has proved to be enough to test the efficiency in each single station, and without introducing any bias in the sample.

3.4.2 Muon trajectory and crossing point determination in the test chamber

The muon crossing point is computed by means of two different approaches. In the first one, valid only for data taken with no magnetic field, the muon trajectory and the subsequent extrapolation to the test chamber is done with a linear fit. The other method relies on the official CMS muon trajectory fitter and propagator in the presence

of a magnetic field. This last method can be used both for runs taken at 0T and 3.8T. A pictorial representation of both extrapolation is shown in Fig. 3.22.

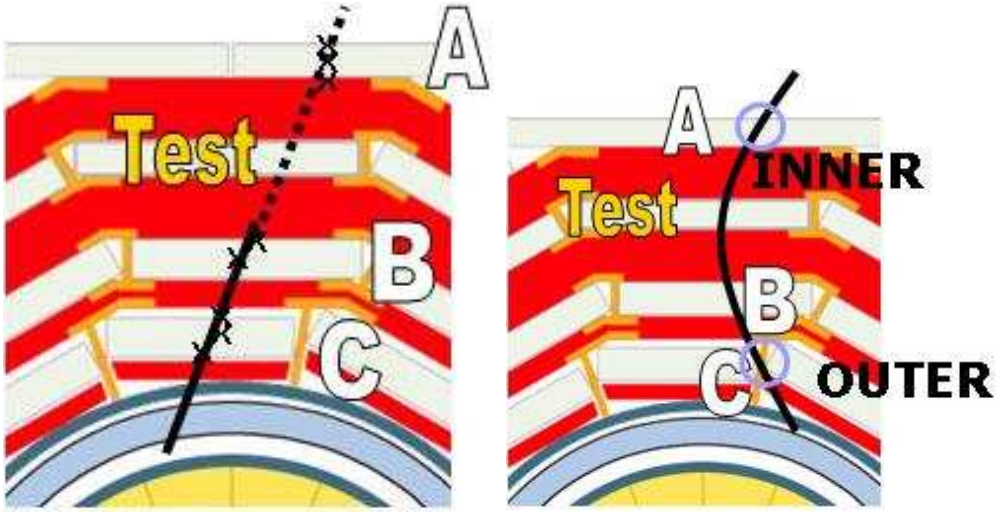

Figure 3.22: Pictorial representation of the reconstruction and extrapolation followed for both 0 T and 3.8 T cases. Chambers A, B and C represent the chambers used to extrapolate to the test chamber.

3.4.2.1 Extrapolation for the 0T runs (linear method)

We have implemented a simple linear extrapolation. Hits in the neighboring chambers were all globally fitted to a unique linear trajectory via a weighted least squares method. A weight was assigned to every hit in order to account for the uncertainty due to multiple scattering, depending on the amount of iron traversed by the muon from that particular hit to the test chamber. This is carried out by assuming a deviation angle from the straight path equal to the scattering angle distribution width. The scattering angle distribution width was taken as (5)

$$\theta = \frac{13.6\text{MeV}}{\beta cp} z \sqrt{\frac{x}{X_0}} \left(1 + 0.038 \ln\left(\frac{x}{X_0}\right)\right) \tag{3.4}$$

We take 20 GeV as the typical momentum for a muon crossing CMS (see Fig. 3.42). In this formula, z is the charge of the incident particle, x the distance traversed and X_0 the radiation length of the material traversed, which is 1.757 cm for iron. The weight is computed as the inverse of the spatial deviation after assuming this scattering angle. In

Table 3.2: Width of the multiple scattering angle distribution for a 20 GeV muon, depending on the amount of iron traversed.

Stations	Distance (cm)	Multiple Scattering Angle Dist. Width, θ (deg)
MB1 - MB2	45.1	0.17
MB3 - MB4	69.68	0.34
MB2 - MB3	70.95	0.35
MB1 - MB3	116.05	0.74
MB2 - MB4	140.63	1.00
MB1 - MB4	185.73	1.53

Tab. 3.2 the scattered angle distribution width, in degrees, depending on the distance traversed by the muon between chambers is shown. A perpendicular incident muon is assumed. Once the fit is performed, the muon crossing point at the test chamber is determined extrapolating the muon path to the chamber half-height. This procedure is followed in both r-ϕ and r-θ planes independently.

In Fig. 3.23 the error in the extrapolated position is shown for each chamber in both planes (only for top sectors). As expected, in Fig. 3.23 (left), the errors for internal chambers (from the extrapolation point of view we call internal those being in the middle of the iron, i.e. MB2 and MB3, and external ones those in the borders, i. e. MB1 and MB4) are smaller, of the order of few millimeters. The errors in MB2 are smaller than in MB3 as a consequence of the iron thickness between MB1 and MB2, being smaller than between MB3 and MB4.

In the r-z plane the picture is different, as only 3 out of the 4 stations contain information in this plane[1]. In this view, Fig. 3.23 (right), the relative values of the errors can also be explained in terms of the number of hits and their positions when performing the fit. In this case, MB2 is the only internal station. As a result, the muon crossing point in this chamber can be extrapolated with the minimal error. Between MB1 and MB3, both external stations, errors are smaller in MB1, as a consequence of the iron width between stations. Errors in the MB4 case are rather difficult to predict. MB4 is the most external station, with a high uncertainty due to the multiple scattering.

[1] 4Dsegments in r-ϕ plane contain up to 8 hits while in r-z view contain up to 4 hits

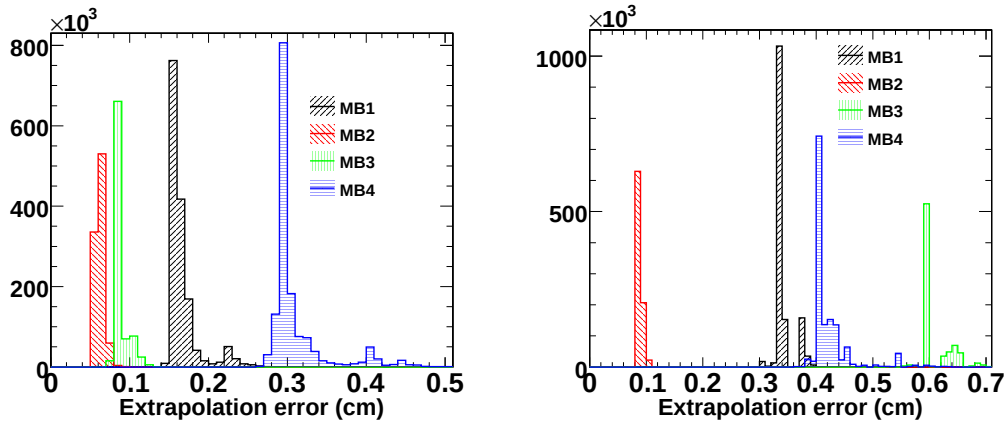

Figure 3.23: Extrapolation error (linear case) - Error in the extrapolated point in the r-ϕ plane (left) and in the r-z plane (right) for all stations in sectors 3, 4 and 5, of all wheels.

But, on the other hand, the number of hits used to perform the extrapolation is higher than in the other stations, resulting in a better extrapolation. As a result of these two opposite effects, the errors for MB4 lie somewhere in between MB1 and MB3.

The extrapolated points can be compared with the actual position of the reconstructed 4Dsegment at the chamber half-height in both views r-ϕ and r-z. If we compute the distance between the reconstructed point (X_{reco}) and the extrapolated one (X_{extrap}) (hereafter $X_{reco} - X_{extrap}$ distribution) in the r-ϕ plane, the result is shown in Fig. 3.24 (left) in this case for chamber W-1 S4 MB1. This distribution is fitted to a gaussian function. The sigma of the gaussian fit (around 3 mm) is higher than the extrapolation error shown in Fig. 3.23 (left). This may be a consequence of the simplifications carried out when applying 3.4. In Fig. 3.24 (right) this distribution is superimposed to those from MB2, MB3 and MB4. The sigma after a gaussian fit of these distributions (0.27 cm, 0.18 cm, 0.21 cm and 0.37 cm for MB1, MB2, MB3 and MB4, respectively) qualitatively agree with the errors from Fig. 3.23 (left).

In Fig. 3.25 the mean value (top) and sigma (bottom) for all MB1 CMS top sector chambers of the gaussian fitted $X_{reco} - X_{extrap}$ distributions in the r-ϕ plane are shown. The mean values are around 0 and not bigger than half a cm. The sigma values are in agreement with the errors previously computed in most of the cases, being around 2 mm in average.

The same has been done in the r-z view. In Fig. 3.26 the $Y_{reco} - Y_{extrap}$ distribution

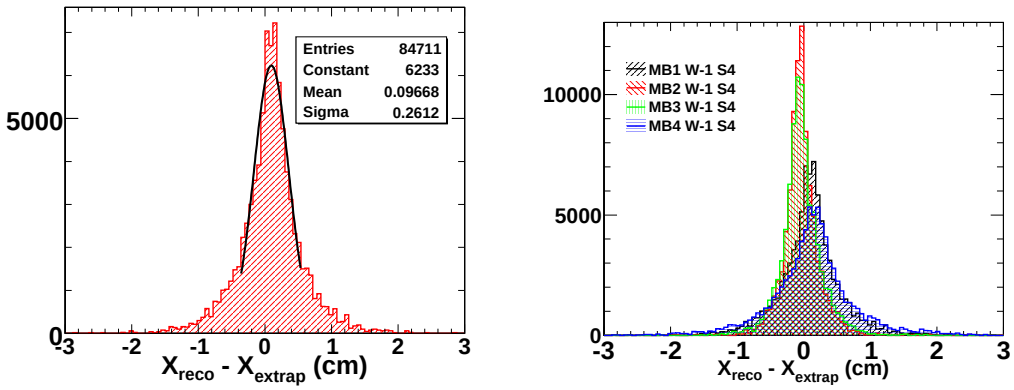

Figure 3.24: Residuals (linear case) - Residual between the extrapolated muon crossing point and the reconstructed 4Dsegment using the linear extrapolation (left) for the W-1 S4 MB1 chamber and (right) for all the stations in W-1 S4, in the r-ϕ plane.

is depicted for chamber W-1 S4 MB1. The sigma of the gaussian fit in the r-z plane is higher than in the r-ϕ one. This is a consequence of the lower number of hits used to build the extrapolation in this case.

In Fig. 3.27 the mean value (top) and sigma (bottom) of the gaussian fitted $Y_{reco} - Y_{extrap}$ distribution for all MB1 CMS top sector chambers are shown. As previously, the average values are around 0 and not bigger than half a cm. The values of the width of the distributions reflect again qualitatively the behavior described regarding the extrapolation error in Fig. 3.23 (right).

3.4.2.2 Extrapolation for the 3.8T runs (STA based method)

For runs taken with magnetic field on, the muon does no longer follow a linear trajectory in the r-ϕ plane. Therefore, we have followed a different approach. We rely on the official methods developed in CMS, and available in CMSSW for muon trajectory fitting and extrapolation (36). The position extrapolated is computed propagating from the track innermost (outermost) position to test MB3 or MB4 in top sectors (MB1 or MB2 in bottom sectors). In this way, the distance from the propagation origin to the test chamber is minimal. One has to notice that in the linear case hits from the test chamber do not take part in the fitting track. However, when computing the fit and the extrapolation using the official code, by means of the so called StandAlone (STA) track, hits from the test chamber are used.

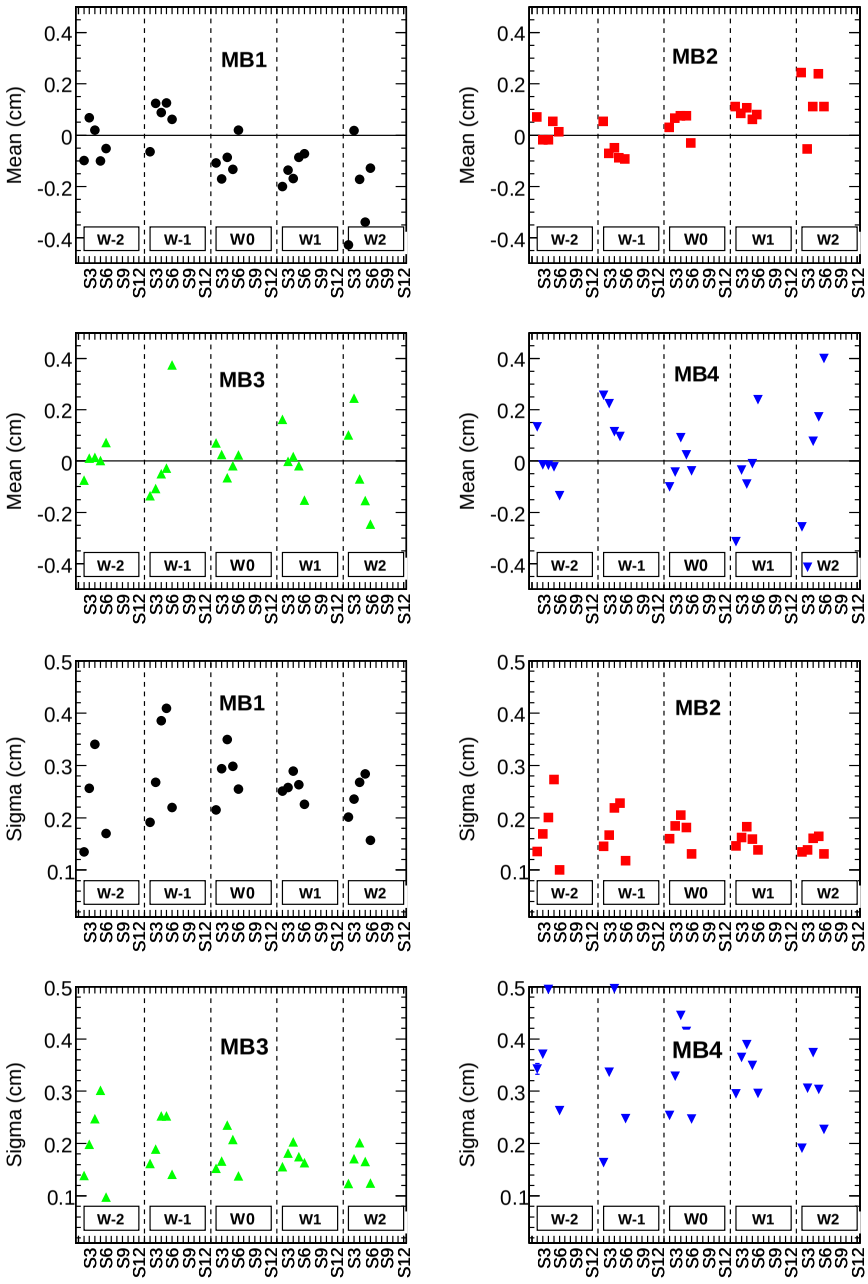

Figure 3.25: Residuals (linear case) - Mean and sigma values of the gaussian fit to the residual between the extrapolated muon crossing point and the reconstructed 4Dsegment (top) mean and (bottom) sigma for all CMS top sector DT chambers, using the linear extrapolation in the r-ϕ plane.

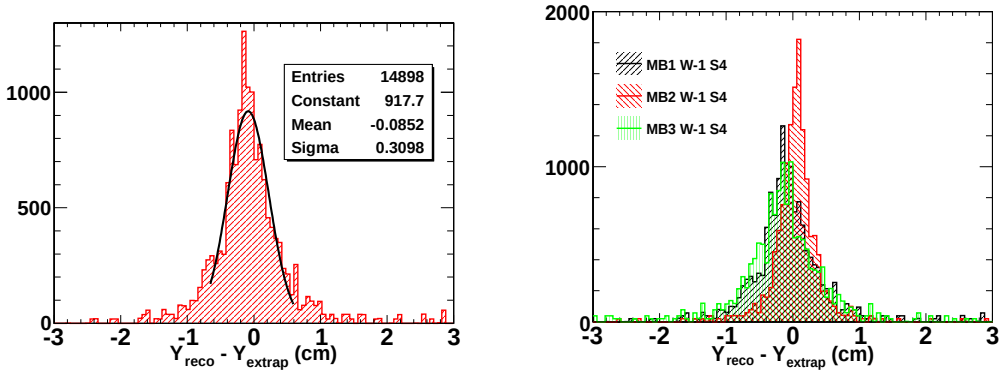

Figure 3.26: Residuals (linear case) - Residual between the extrapolated muon crossing point and the reconstructed 4Dsegment using the linear extrapolation (left) for the W-1 S4 MB1 chamber and (right) for all the stations in W-1 S4, in the r-z plane.

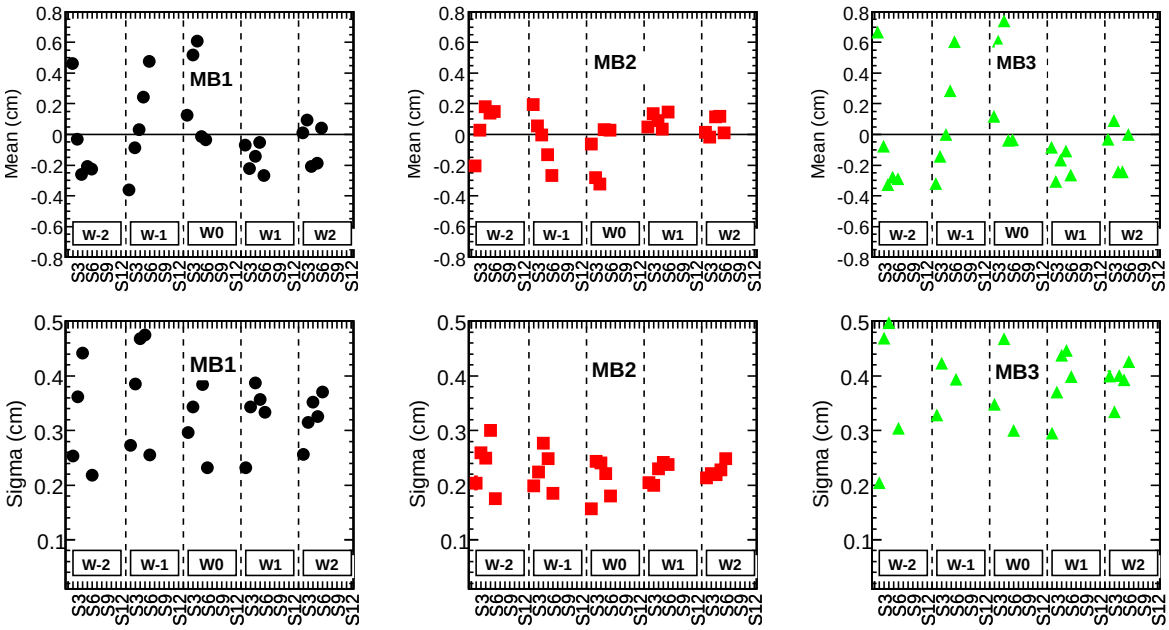

Figure 3.27: Residuals (linear case) - Mean and sigma of the gaussian fit to the residual between the extrapolated muon crossing point and the reconstructed 4Dsegment (top) mean and (bottom) sigma for all CMS top sector DT chambers using the linear extrapolation in the r-z plane.

In Fig. 3.28 (left) the extrapolation error for all MB1, MB2, MB3 and MB4 stations in top sectors in r-ϕ is shown. The propagation error is minimal for MB1 and MB4 (as the distance to propagate from the innermost or outermost position is minimal for them), and maximal for MB3, where the propagating distance is maximal.

In Fig. 3.29 (left) the difference between the extrapolated local coordinate and the reconstructed one in the r-ϕ plane is depicted for W-1 S4 MB1. In Fig. 3.29 (right) the same quantity is plotted for W-1 S4 MB1, MB2, MB3 and MB4. The ordering of the width of these distributions is in agreement with the ordering of the propagation error shown before. Compared with the previous linear method, a higher accuracy in the position determination is reached, since the residual distribution is shown to be narrower. One has to remember that in the STA case an optimized fit procedure is established (a successive Kalman filter procedure going back and forth in the track) [REF], and applied several times, and hits in the test chamber are considered in the STA method, consequently reaching better results.

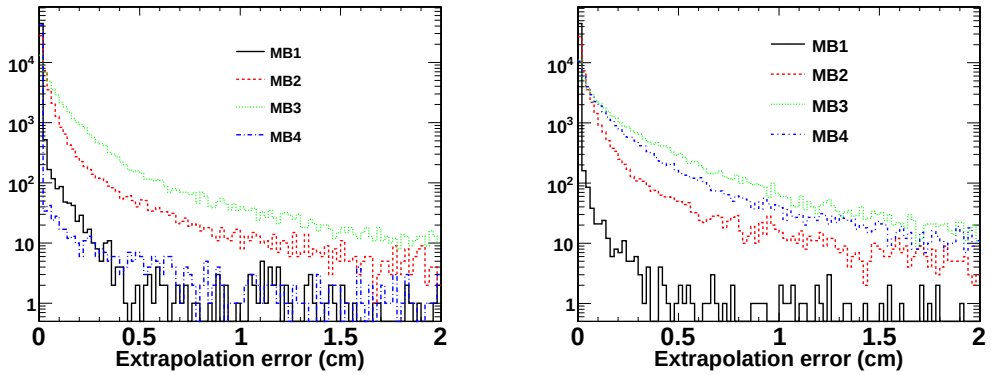

Figure 3.28: Extrapolation error (3.8 T run case) - Extrapolation error for MB1 (black solid line), MB2 (red dashed line), MB3 (green dotted line) and MB4 (blue dashed-dotted line) in (left) r-ϕ and (right) r-z planes.

In Fig. 3.30 the mean value (top) and sigma (bottom) of the gaussian fitted to the $X_{reco} - X_{extrap}$ distribution for all CMS top sector chambers are shown, now for the STA track extrapolation. The widths in this case are of the order of half the widths in the linear case.

In the r-θ plane the situation is very similar. In Fig. 3.28(right) the extrapolation error in this plane is shown for all the top sectors MB1, MB2, MB3 and MB4. The

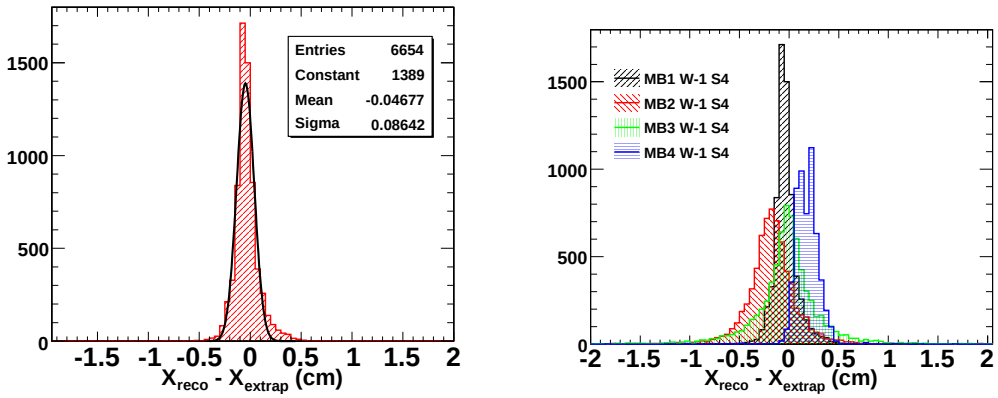

Figure 3.29: Residual (3.8 T run case) - Residual between the extrapolated muon crossing point and the reconstructed 4Dsegment using the StandAlone fitter (left) for the W-1 S4 MB1 and (right) for all the stations in W-1 S4 in the r-ϕ plane.

distribution $Y_{reco} - Y_{extrap}$ is also shown (but now for the r-z plane), in Fig. 3.31, for W-1 S4 MB1 (left) and all the stations in W-1 S4 (right). This distribution is wider than in the r-ϕ case, as expected, since the number of hits per fit is smaller. In Fig. 3.32 the mean value (top) and sigma (bottom) of the previous distributions are shown for all CMS top sector chambers, now for the STA track extrapolation.

3.4.3 Quality criteria

The necessary conditions to guarantee a robust efficiency determination are presented in this section. This efficiency is defined as a ratio of cases where a particular DT chamber reconstructs a 4Dsegment over the total possible number of cases where a muon traverses this chamber. Based on the results presented in the previous section, a maximum residual value of 5 cm between the reconstructed position and the extrapolated one is allowed, in both planes and both magnetic field configurations. All chambers with a 4DSegment reconstructed beyond this 5 cm around the extrapolated coordinate will be considered as non-efficient ones. In this way, there are three possible sources of inefficiency: lack of a 2D segment (one of the r-ϕ or r-z views) inside the 4D segment, lack of the 4D segment itself or a distance between extrapolated and reconstructed position above the threshold. Since we are interested in measuring the DT chamber efficiency, not its accuracy, this threshold of 5 cm comes out to be a rather loose re-

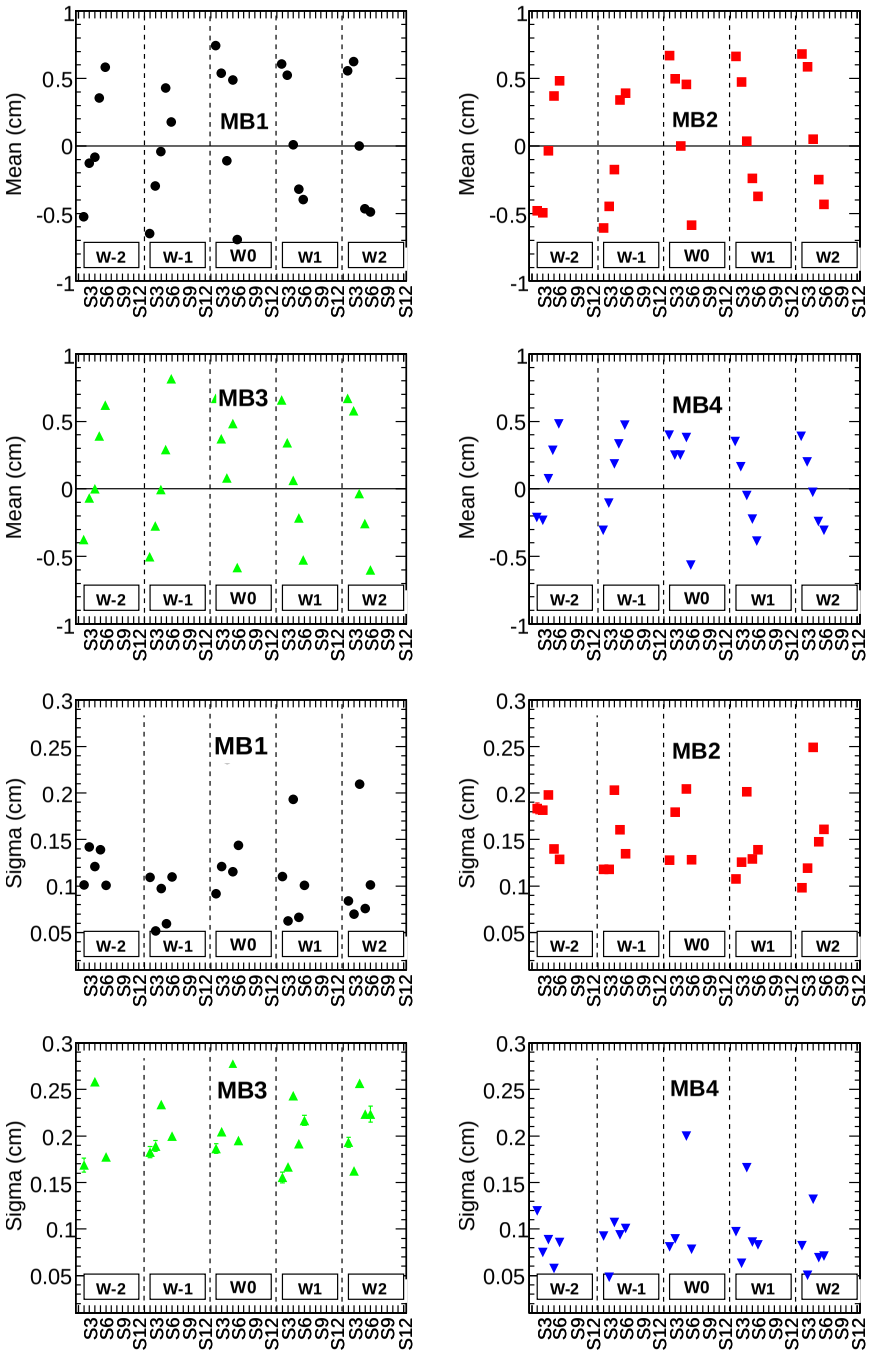

Figure 3.30: Residual (3.8 T run case) - Mean and sigma of the gaussian fit to the residual between the extrapolated muon crossing point and the reconstructed 4Dsegment (top) mean and (bottom) sigma for all CMS top sector DT chambers using the StandAlone fitter and the SteppingHelix propagator in the r-ϕ plane.

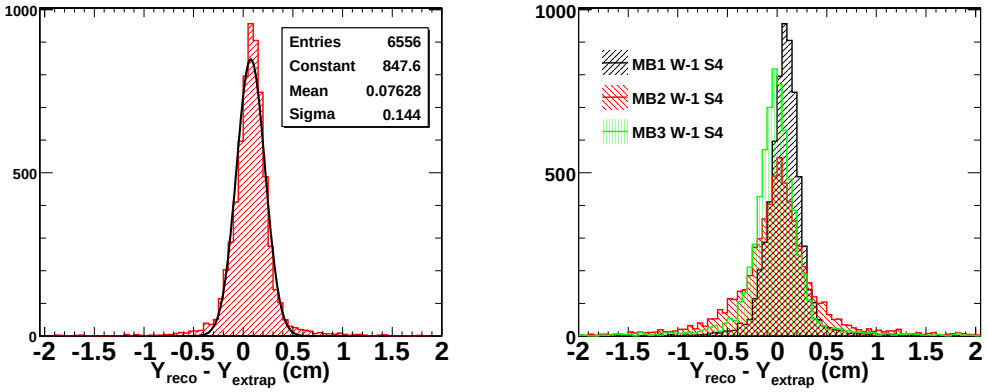

Figure 3.31: Residual (3.8 T run case) - Residual between the extrapolated muon crossing point and the reconstructed 4Dsegment using the StandAlone fitter and the SteppingHelix propagator (left) for the W-1 S4 MB1 and (right) for all the stations in W-1 S4 in the r-z plane.

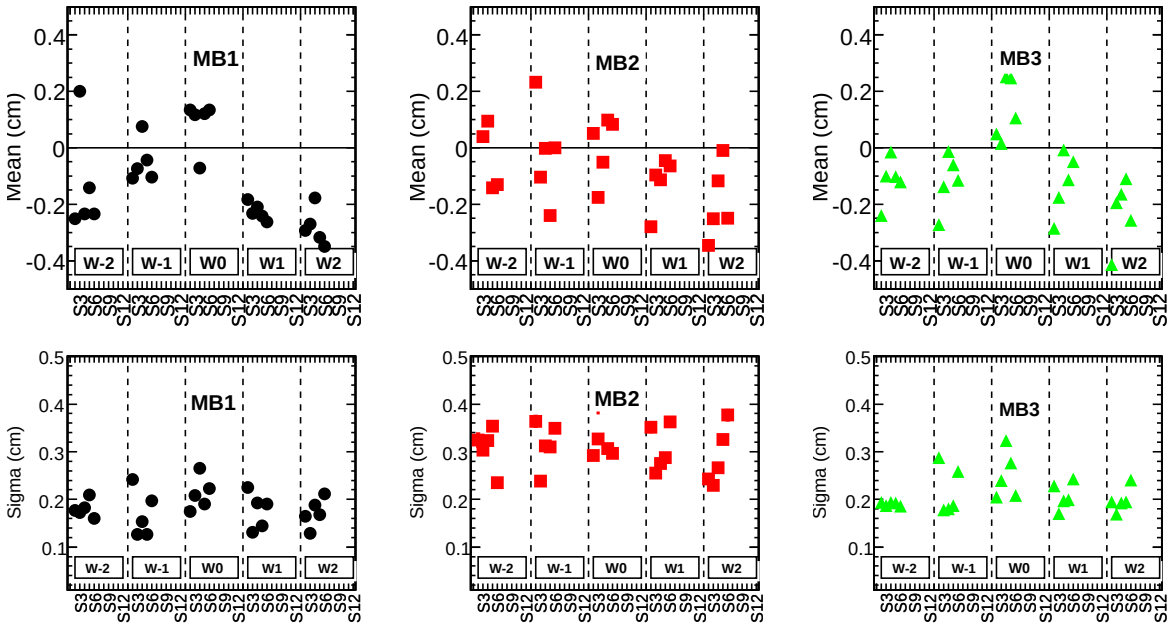

Figure 3.32: Residual (3.8 T run case) - Mean and sigma of the gaussian fit to the residual between the extrapolated muon crossing point and the reconstructed 4Dsegment: (top plot) mean and (bottom plot) sigma for all CMS top sector DT chambers using the StandAlone fitter and the SteppingHelix propagator in the r-z plane.

quirement (around 20 sigmas of the distance extrapolated to reconstructed position distribution).

The efficiency values to be computed are essentially 100%, finding only very few cases where the chamber is not efficient. If, by any reason, the extrapolation is wrongly determined and gives rise to requiring a (non-existent) 4Dsegment to be present in the chamber, the error induced in the efficiency calculation is sizeable. In order to guarantee good quality extrapolations, a cut in the χ^2/ndof of the linear fit is required to be less than 1 in both planes, for the linear method . In Fig. 3.33 (left)a histogram of the χ^2/ndof of the linear fit is depicted for both r-z and r-ϕ planes for the linear method. Setting a maximum value of 1 is shown to be suitable for this case and not really restrictive, as most of the sample fulfils this requirement. This cut in the $\chi^2/ndof$ is rather loose but it allows to remove pathological cases.

In addition to this cut, in order to ensure a good precision in the extrapolated position, a lower cut in the minimum number of hits that have been used to perform the extrapolation is applied. In Fig. 3.33 (right) the number of hits used to perform the extrapolation is shown for the STA method in both the r-ϕ and r-z planes, where the number of hits from the test chamber are not included in the *hit-counting*. In Fig. 3.34 the efficiency versus the number of hits in the r-ϕ plane of the STA track is shown for the STA based method. We can appreciate that for more than 13 hits (the minimum number of hits set) the efficiency reaches a plateau for all the stations, therefore a cut at 14 (9) hits is set for the STA (linear) method respectively, not taking into account the number of hits from the test chamber. In the r-z plane only tracks with at least 5 hits, apart from those in the test chamber, have been considered in both methods. Requiring a minimum of 5 hits means, in practice, imposing the existence of at least two chambers to perform the extrapolation.

3.4.4 Inefficiencies due to geometrical effects

The efficiency is computed in two steps. First we will show the efficiency computation for the top sectors (all the sectors placed in the top hemisphere of CMS: S2, S3, S4, S5, and S6). Then we will see the particularities for the bottom sectors (sectors placed in the bottom hemisphere of CMS: S8, S9, S10, S11 and S12). S1 and S7 are not considered in this study. Being vertical sectors, the amount of cosmic muons reconstructed is not enough to perform a complete study on the reconstruction efficiency.

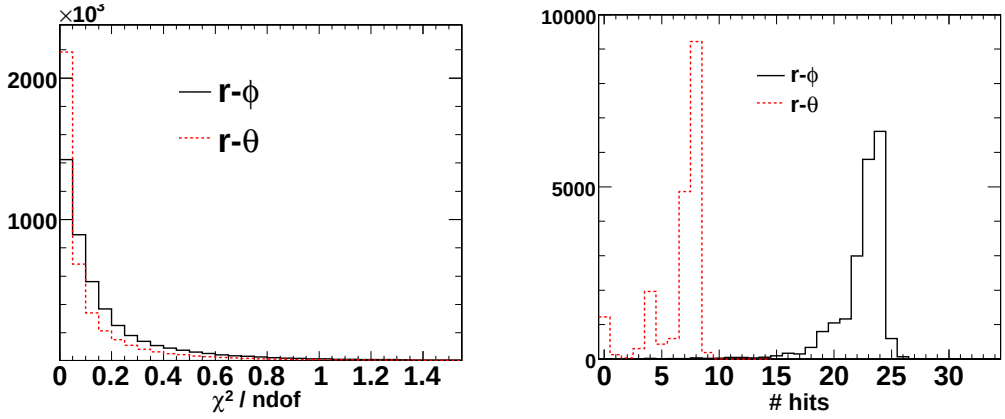

Figure 3.33: Quality criteria - χ^2/ndof of the linear fit for both r-θ (red dashed line) and r-ϕ (black solid line) planes using the linear extrapolation (left) and number of hits in the STA track (not considering the number of hits in the test chamber) in r-ϕ (black solid line) and r-θ (red dashed line) planes (right).

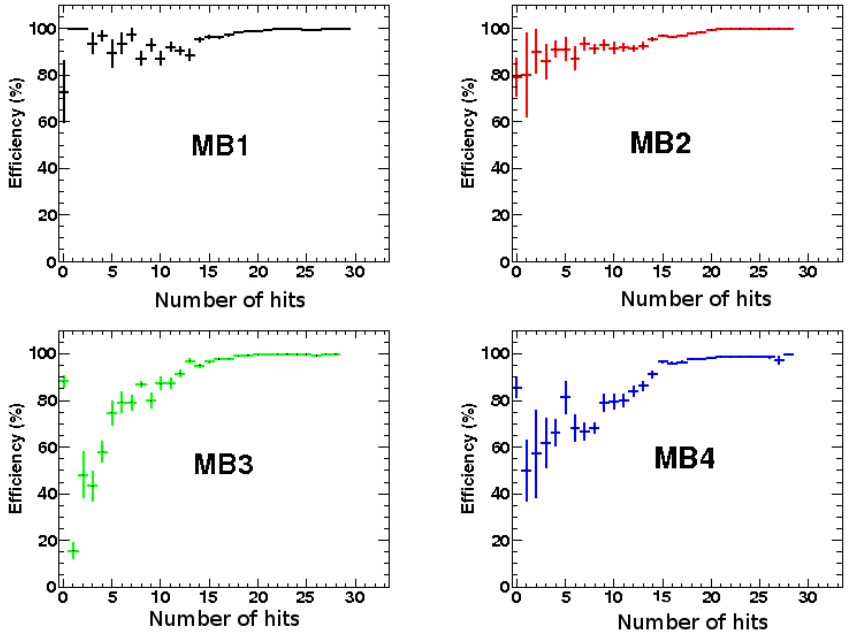

Figure 3.34: Efficiency versus the number of hits - DT 4DSegment efficiency versus the number of hits in the r-ϕ view of the STA track (no hits from the test chamber are considered in the *hit-counting*) for MB1 (black), MB2 (red), MB3 (green) and MB4 (blue).

The previously defined 4DSegment efficiency is computed over the resulting good-quality muon sample. The results are shown in Fig. 3.35, where the efficiency for all CMS top sector stations is plotted. This figure shows that efficiencies are high (more than 90% almost everywhere), but also some other unexpected effects are revealed. Some differences between stations (e.g. MB3 being more efficient than MB1) can be appreciated. Also, within the same station, a certain pattern can be observed, depending on the sector, which is repeated for every wheel. Since the chamber construction and installation process is the same for all stations and sectors, no intrinsic differences among efficiencies are expected.

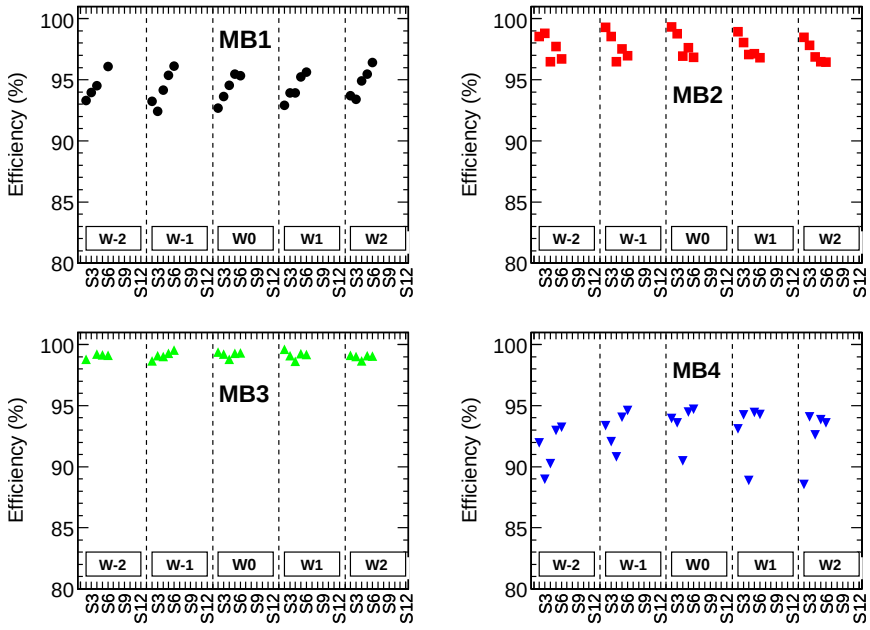

Figure 3.35: Efficiency - 4DSegment reconstruction efficiency for all CMS top sector chambers for a 3.8 T run.

In general, these drops in the efficiency are due to cases where the muon crosses the test station near the borders. Due to the extrapolation error, a 4Dsegment may be looked for in a chamber border when the muon has not actually gone through it. In Fig. 3.36 the DTSegment efficiency versus the chamber position of the extrapolated crossing point is depicted, where the drop in the efficiency near the chamber borders is clearly seen. The non-sensitive part between the first cell and the physical limit of the

chamber (which should not be considered as an intrinsic inefficiency but of a geometrical nature) and also some occasional electrical noise may add to this inefficiency.

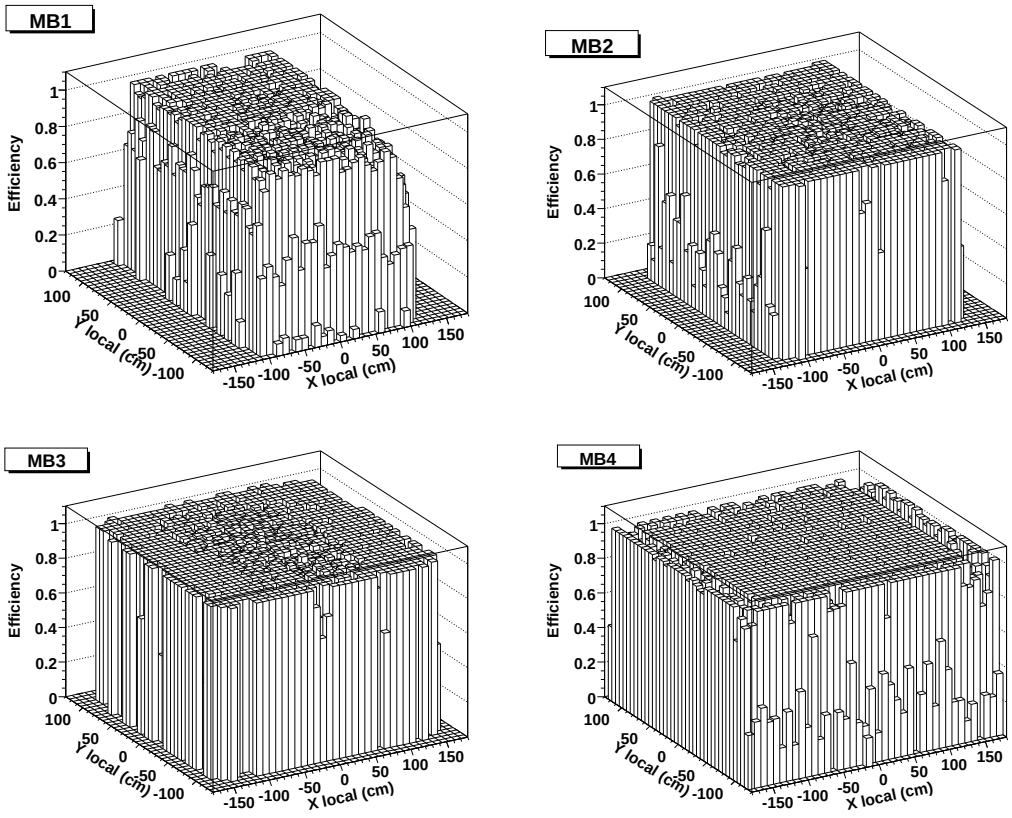

Figure 3.36: Efficiency versus the chamber position - Efficiency versus the muon crossing point position in the test chamber for (top-left) MB1, (top-right) MB2, (bottom-left) MB3 and (bottom-right) MB4 using the STA based method.

The observed efficiency differences among stations in Fig. 3.35 can be explained as a combination of this border effect, the requirements imposed in the sample selection and the chamber geometrical disposition in CMS (see Fig.3.4). The requirement of having three other 4Dsegments in the same wheel and sector affects unavoidably the muon crossing point spatial distribution in the test chamber. Since MB2 and MB3 are both internal chambers, and all the four stations have the same length in the z coordinate, it is unlikely to have a muon crossing MB2 or MB3 z border once the previous requirement is imposed. Something similar happens in the r-ϕ plane border

region. In the selected sample it is more likely to have muons crossing the border of the chambers in case of the external ones. In particular, the high size of MB3 in the r-ϕ plane "protects" this station against the border effect. The different magnitude in which the border effects are present in each station can be directly seen in Fig. 3.36. The extrapolated position for muons in which the test chamber is not efficient is shown in Fig. 3.37. It is interesting to point out from this figure the efficiency drop due to the physical gap between MB4 S4 and MB4 S13. As a concluding remark, geometrical border effects affect the total efficiency with a value that varies from 5% inefficiency in MB4 S4 to the 0.1% in MB3 S4. Taking into account these considerations, the differences in efficiency between MB3 stations on one side, and MB4 and MB1 ones (the external stations) can be understood.

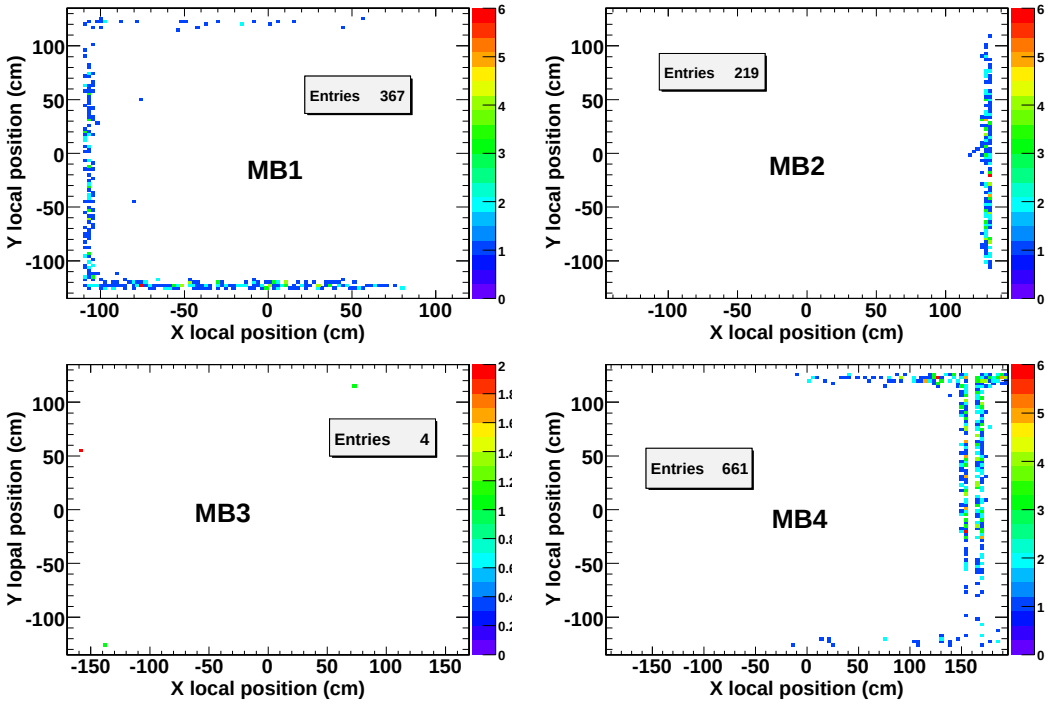

Figure 3.37: Inefficiency - Extrapolated position in the test chamber for those muons where the test station fails to reconstruct a 4Dsegment for MB1 (top-left), MB2 (top-right), MB3 (bottom-left) and MB4 (bottom-right) in a 3.8T run.

Concerning the efficiency pattern correlated with the sector number, seen in Fig. 3.35, it is, again, due to the combination of the border effect, the requirements imposed in

the sample selection and the chamber disposition in CMS, not designed to receive cosmic muons, but muons from the interaction point. In particular, the last two elements make the muon impact position distribution in a station to be rather different depending on the sector considered. As a result, not every sector will be affected by the border effect in the same magnitude. In Fig. 3.38 the muon crossing position distribution in the test chamber, as given by the extrapolation, is shown for MB1 W-1 in sectors 3, 4 and 5. The population of muons arriving close to the border is higher for sector 3 and therefore the border effect in this sector is more relevant than in sector 5.

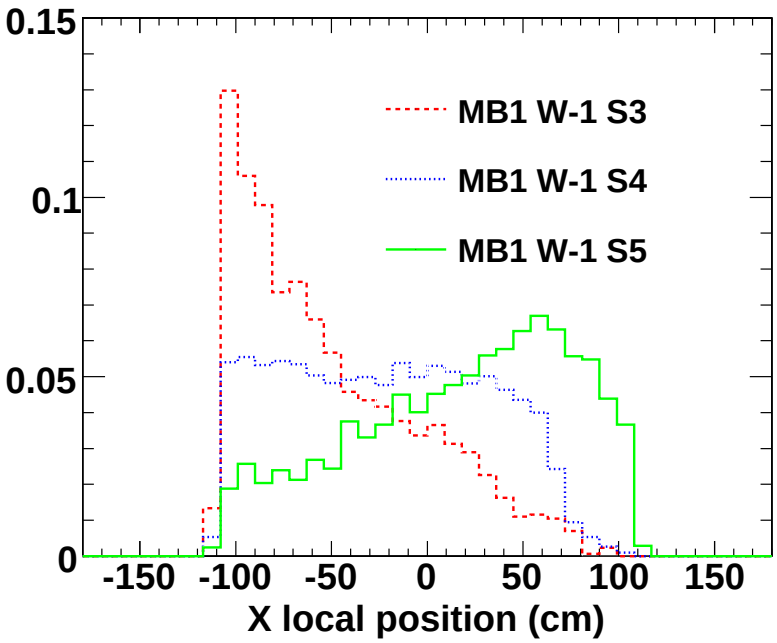

Figure 3.38: Extrapolated position - Extrapolated local x position distribution of the incident muon in MB1 station Wheel -1, for sectors S3 (red solid line), S4 (blue dashed line), S5 (green dotted line).

As we are interested in an intrinsic efficiency computation, a fiducial volume is defined in each test chamber, maximizing its volume to include the whole test station except a frame of 8 cm width at the borders. It is important to remark that in proton-proton collision none of this considerations are necessary, since all the muons come from the interaction point, from which all sectors are identical due to rotational symmetry of the CMS detector.

In the next section, the 4D-Segment reconstruction efficiency in CMS DT chambers will be quantified, inside this fiducial volume. For that purpose, the muons where the 4Dsegment is reconstructed in the test chamber within a certain margin around the extrapolated position are computed. It will be applied to both r-ϕ and r-z planes and then, to a combination of both.

3.4.4.1 Results for top sectors

As a first step, the efficiency is computed for CMS top sectors. In Fig. 3.39 the efficiency for all CMS top sectors in the r-ϕ plane (top) and r-z plane (bottom) is shown for MB1, MB2, MB3 and MB4 stations for the run taken at 0T, and using the linear method. In Fig. 3.40 the same quantities are plotted for the STA method, for a 3.8T run. Comparing both methods we can appreciate the compatibility between both results, and the high efficiency values for most stations (over 97% in the r-ϕ view). The r-θ view presents smaller values for the efficiency (2%-3% lower values) related to the smaller number of hits in this view, and thus, the less precise extrapolation. Moreover, the smaller number of hits per station in the r-θ view, 4 hits, compared with the 8 hits in the r-ϕ plane makes the loss of a θ segment more likely than the loss of a ϕ segment.

3.4.5 Results for bottom sectors

The efficiency for bottom sectors for a 3.8T run, using the same analysis procedure, is shown in Fig. 3.41. The bottom sectors efficiency appears to be lower in the MB3 station, when compared to top sectors. Also the efficiency depends on the sector studied (being lower for S8 and S12).

The origin of this efficiency drop lies in the lower transverse momenta distribution of muons reaching the bottom sectors when compared to that in the top sectors. In Fig. 3.42, the muon transverse momentum distribution is plotted for MB1, MB2, MB3 and MB4 stations for sector 4 (left) in the upper hemisphere and sector 10 (right) in the lower one. A shift towards lower momentum values can be observed when going from MB4 S4 to MB1 S4 and then from MB1 S10 to MB4 S10. This shift corresponds to the muon energy loss when traversing CMS. For instance, in the MB3 case, sector 4, only 0.7% of muons have transverse momenta below 5 GeV/c, while up to 8.4% of muons going through MB3, sector 10, have less than 5 GeV/c. As most of the muons follow a path top-down in the detector, it is expected to have lower transverse momentum

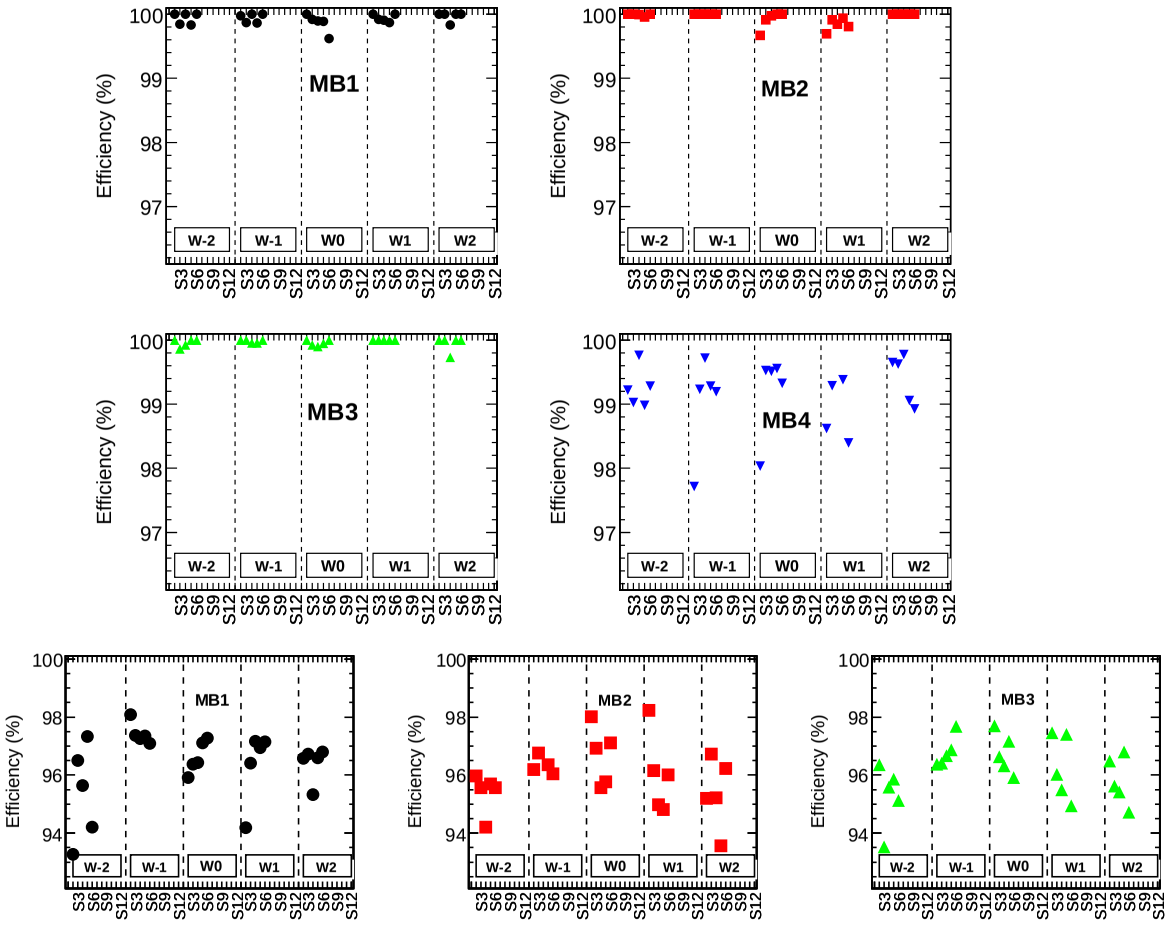

Figure 3.39: Efficiency (linear case) - Efficiency for all DT chambers in CMS top sectors in (top) r-ϕ and (bottom) r-θ planes for the linear method (no magnetic field run).

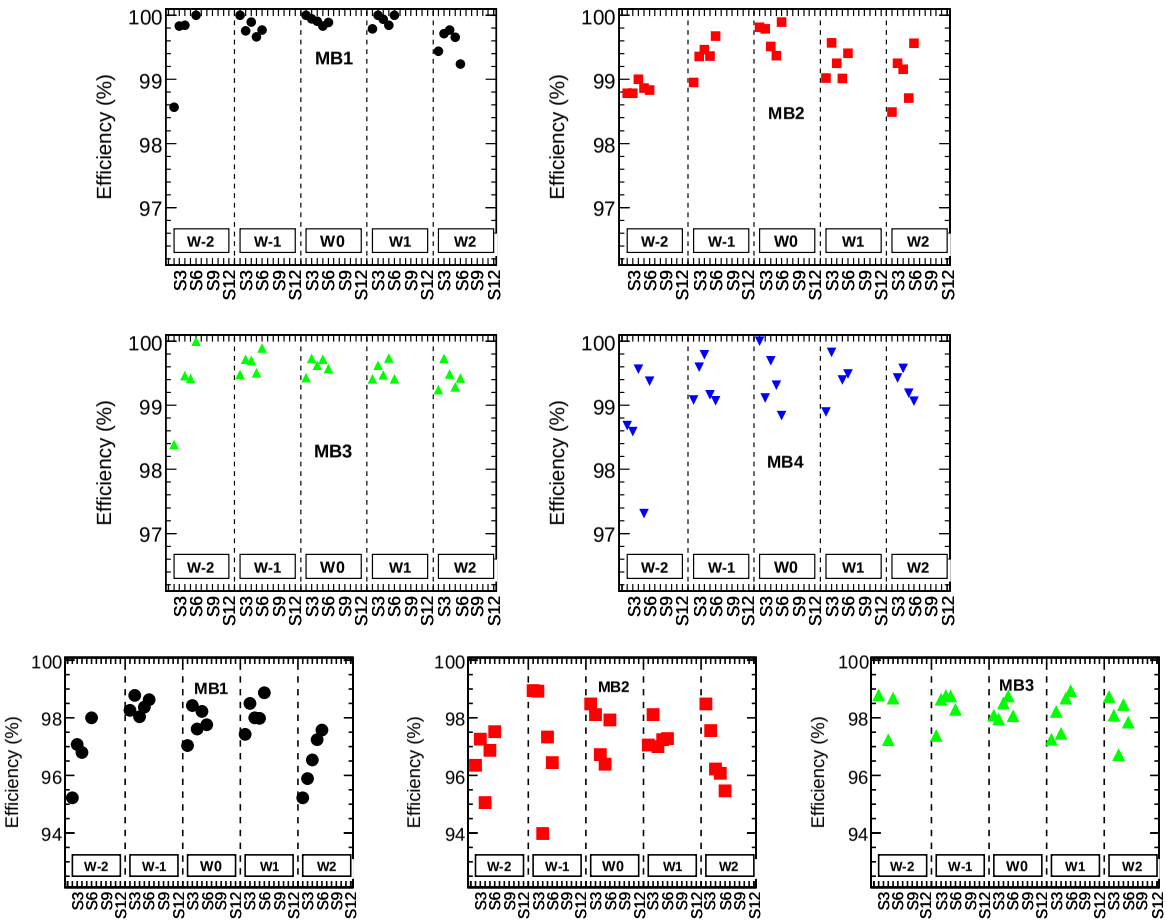

Figure 3.40: Efficiency (3.8 T run case) - Efficiency for all DT chambers in CMS top sectors in (top) r-ϕ and (bottom) r-θ planes for the STA based method (3.8T run).

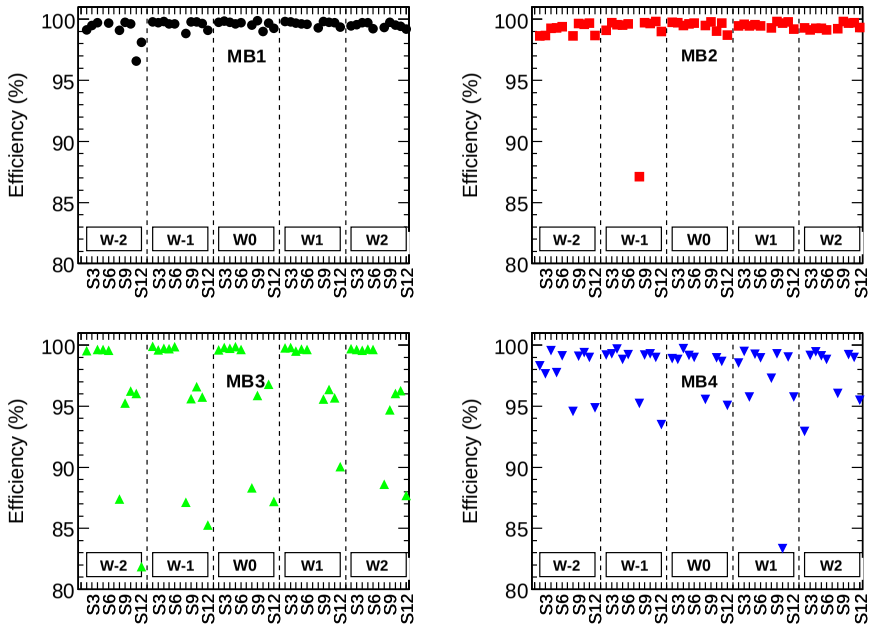

Figure 3.41: Efficiency - Efficiency for all DT chambers in CMS top and bottom sectors in the r-ϕ plane for the STA based method(3.8T run).

distributions in the CMS bottom part. The unlikely low transverse momentum muons (pt less than 5 GeV/c) in the top sectors, turn into more numerous in the bottom ones.

Low momentum muons undergo higher multiple scattering when going through the CMS barrel iron, being more difficult to extrapolate. In Fig. 3.43, the propagation error in the r-ϕ plane versus the muon transverse momentum is shown for W-1 S10 MB3. The plot on the right is an inset of the small extrapolation error region. It can be concluded that low transverse momentum muons present higher propagation errors. In Fig. 3.44 the 4D-Segment reconstruction efficiency is plotted as a function of muon transverse momentum for MB2 and MB3.

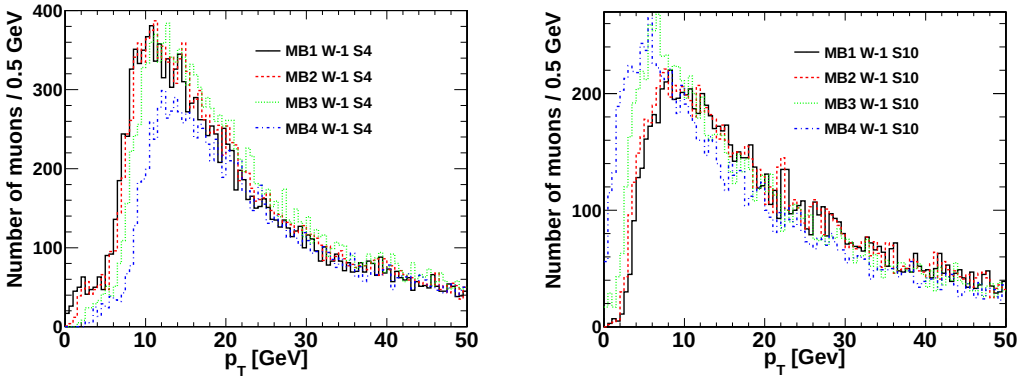

Figure 3.42: Muon transverse momentum distribution - Muon transverse momentum distributions for MB1 (black solid line), MB2 (red dashed line), MB3 (green dotted line) and MB4 (blue dashed-dotted line) in (left) an upper sector (S4) and (right) a lower sector (S10).

The muon propagation through the barrel chambers is performed such that it starts either from the innermost station (MB1) or the outermost one (MB4). In the case of computing the extrapolated position in any of these two stations (MB1 and MB4) the propagation is minimal, taking place inside the chamber, and thus the propagation error is also small. In the case of the inner stations MB2 and, mostly, MB3 the amount of iron traversed in the propagation is maximal, so being the propagation error. In these terms, it is comprehensible to have a wider distribution of the difference between the extrapolation point and the reconstructed one, leading to a higher rate of cases where the reconstructed segment lies beyond the 5 cm window imposed. Within these conditions, it is reasonable that efficiency (in the way it has been defined) in these

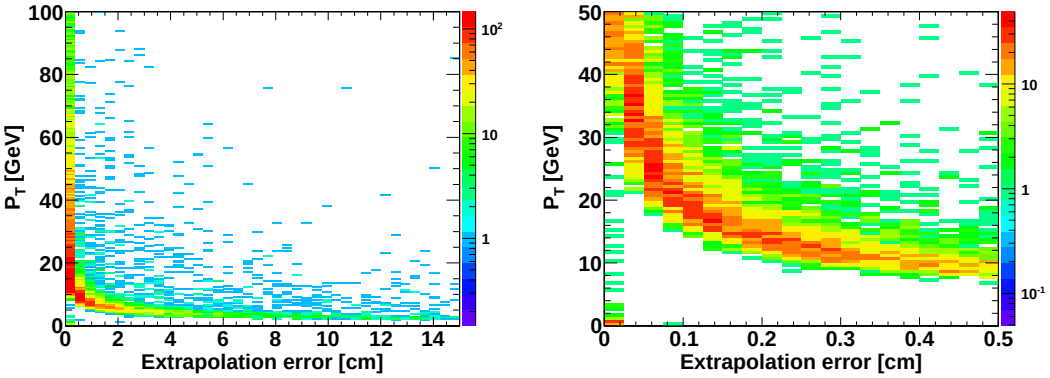

Figure 3.43: Extrapolation error versus muon p_T - Extrapolation error versus the incoming muon transverse momentum in the W-1 S10 MB3. The right plot is a zoomed version of the left one, in the low extrapolation error region.

chambers for low momentum muons is lower. MB1 and MB4, chambers which do not require any propagation to obtain the extrapolated position, are not sensitive to muon momentum.

Concerning the differences in the efficiency values among sectors for the MB3 station, shown in Fig. 3.41, the situation is illustrated in Fig. 3.45. Here the muon momentum distribution in MB3 for different bottom sectors is plotted. The differences between S8 and S12 with respect to S9, S10 and S11 can be easily appreciated. Low momentum muons, when entering CMS vertically from MB4 S4, are bent by CMS solenoid in many cases towards S12 or S8. This fact makes these sectors to have a remarkable amount of low momentum (below 5 GeV/c) muons. Consequently, these sectors will undergo the previously analyzed effect (high extrapolation error) in a bigger proportion.

To prevent this effect, a maximum value of the propagation error is set at 1.5 cm for muons entering the efficiency calculation, which ensures a good extrapolation point.

Fig. 3.46 shows the final efficiencies for all CMS MB1, MB2, MB3 and MB4 stations for the r-ϕ view. Fig. 3.47 present the efficiencies now for the r-z view. In these plots, data from the 4 runs in study have been used (67818, 67838, 68021 and 68141). The same conclusions pointed out before can be also drawn when observing these plots: high efficiency in all CMS chambers (above 99% in r-ϕ), 2% lower efficiencies in r-z and high uniformity in the efficiency values for different sectors and wheels.

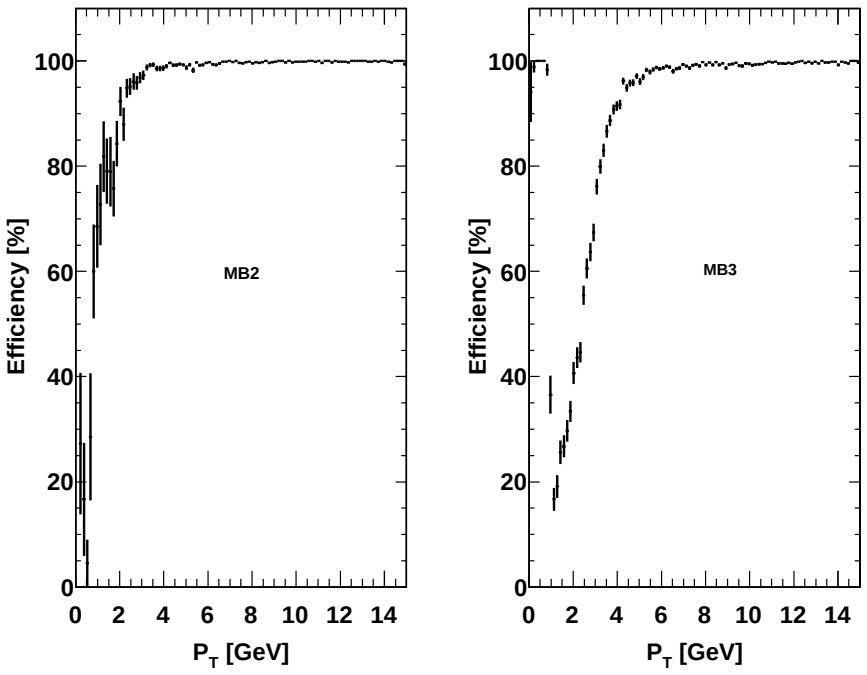

Figure 3.44: Efficiency vs $\mathbf{p}_T$ - DT 4D-Segment reconstruction efficiency versus the incoming muon transverse momentum for (left) MB2 and (right) MB3 in S10.

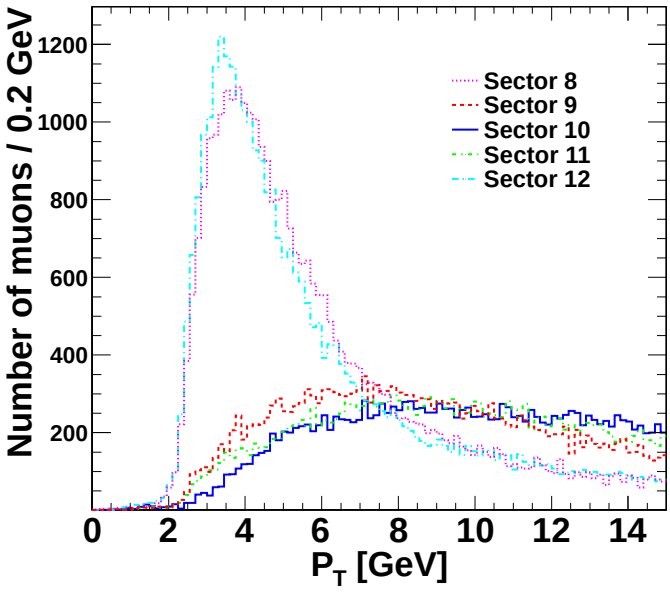

Figure 3.45: Muon p_T distribution - Muon transverse momentum distribution in MB3 station for S8 (blue dotted line), S9 (red dashed line), S10 (dark blue solid line), S11 (green dotted-short dashed line) and S12 (purple dotted-long dashed line).

In Fig. 3.48 a bi-dimensional plot showing the combined 4D-Segment efficiency for all sectors (except the vertical ones, S1 and S7) and all wheels and stations is shown. This efficiency is determined as a combination of both r-ϕ and r-z efficiencies. Low efficiency values can be spotted quickly with the help of the colour code, looking for detector problems or features (e.g. chambers partially off). In the particular case shown here, the MB4 chamber in sector 9 wheel 2 looks inefficient. Fig. 3.49 shows the number of hits in each cell (occupancy) of every chamber in a given sector and wheel (in this case sector 9 and wheel 2). It is shown that MB4 was inactive during part of the run, giving rise to the inefficiency observed in the previous plot.

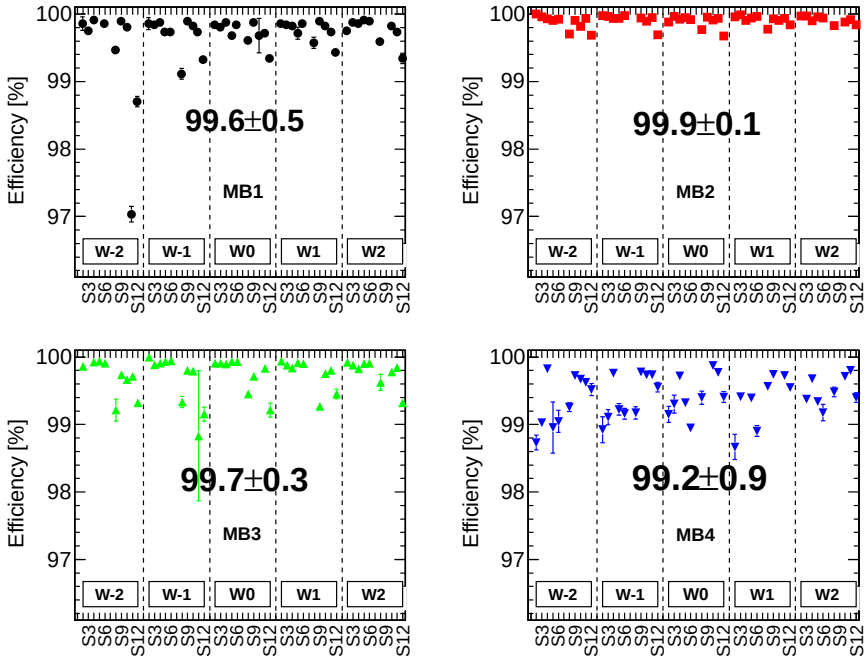

Figure 3.46: Final efficiency - CMS DT 4DSegment reconstruction efficiency in the r-ϕ plane.

3.5 Validity tests

When computing the efficiency with the STA based method, a potential bias arises since hits from the test chamber are included in the STA track fit procedure. This is not the case in the linear method where hits from the test chamber are explicitly

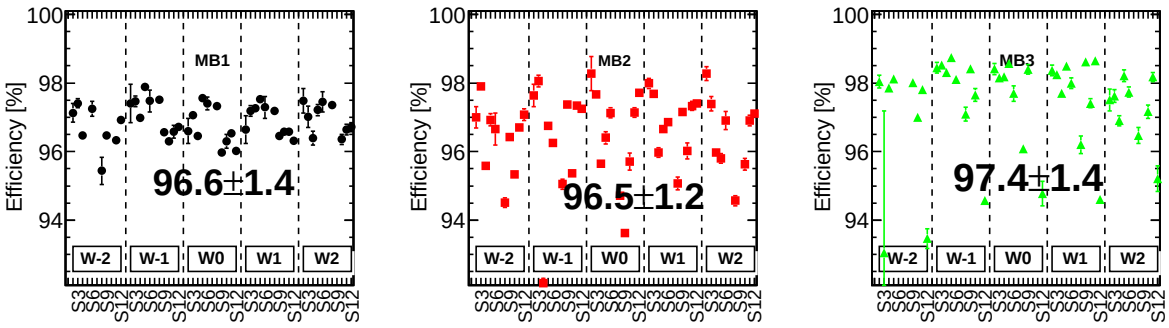

Figure 3.47: Final efficiency - CMS DT 4DSegment reconstruction efficiency in the r-θ plane.

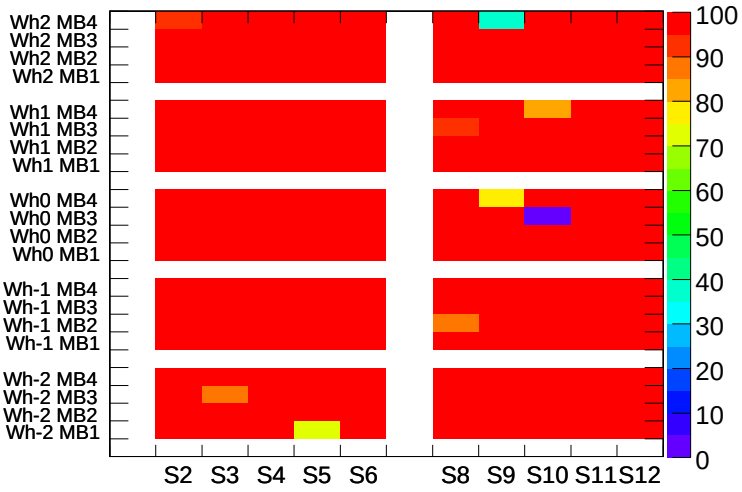

Figure 3.48: Final efficiency - Combined r-ϕ and r-z CMS DT 4DSegment reconstruction efficiency.

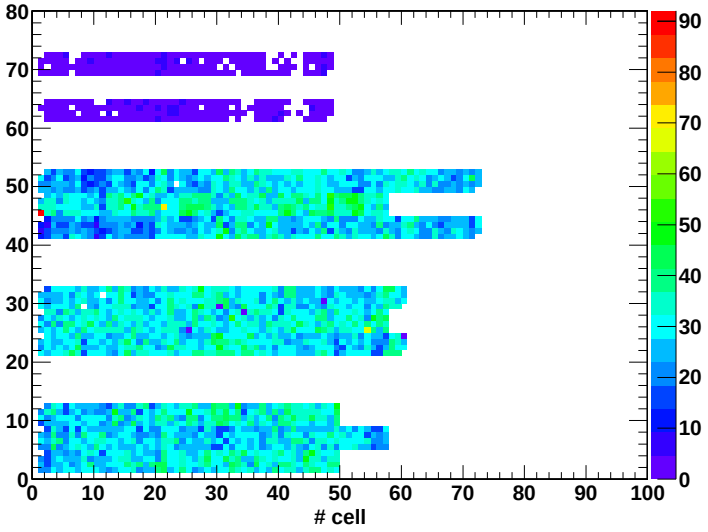

Figure 3.49: Occupancy distribution - Occupancy distribution for Wheel 2 Sector 9. MB4 chamber was inactive during part of the run (low occupancy).

excluded from the fit and extrapolation. A possible bias may arise in two different ways: first when performing the sample selection, second when propagating to the test chamber from the inner or outermost position.

When producing the STA track, hits from the test chamber may help the Producer to build the track. Efficient cases will then be in higher proportion, biasing the sample. However, the restrictive cut set in the minimum number of hits in the STA track, in addition to the requirement of having a 4Dsegment in each of the other 3 chambers in the same wheel and sector as the one being tested, allows only good tracks to be computed. These two requirements guarantee that in most cases the STA producer does not depend on hits in the test chamber. For this clean sample the STA is produced independently whether a 4Dsegment in the test chamber exists or not. In order to check this assumption eight 3.8T runs, 4 runs in which one of the stations was inactive (MB2 W-2 S11) and 4 ones in which all the 4 stations in W-2 S11 were active, were analyzed (see Tab. 3.3).

If the study is biased, the proportion of events after the sample selection should be higher for runs having 4 active chambers (as a consequence of the test chamber contribution to the STA producer). However, this is not what it is observed. Fig. 3.50

shows the proportion of events with at least one STA track in two cases: before any requirement and after asking for a minimum number of hits and a maximum propagation error, for the 8 runs. In the first case there is a bias effect of around 1% difference bewteen both kind of runs. This can be understood since the test chamber is helping the STA producer to build the STA track.

However, in the second case, the quality cuts are somehow allowing only good tracks to become part of the selected sample, thus rejecting those cases in which the test chamber is essential to build the STA track. As seen in this figure, the bias introduced in this study by the use of the STA track is more conceptual than practical. In order to crosscheck this conclusions, a 0T run has been analyzed independently, with the non-biased linear method and with the STA based method. Results in both cases are identical.

The second possible bias source in the study is the contribution of the test chamber hits to the extrapolation. These hits will make the residual distribution to be narrower, but will not make inefficient cases become efficient. Bad reconstructed segments (e.g. with a different slope with respect to the STA track) will not take part in the STA track, thus avoiding any kind of bias in the extrapolation.

Table 3.3: Runs used to test the possible bias in the study.

Run	**Active stations in W-2 S11**	**Total number of events analyzed (Mev)**
67818	MB1-MB2-MB3-MB4	1.9
67838	MB1-MB2-MB3-MB4	1.8
68021	MB1-MB2-MB3-MB4	1.8
68141	MB1-MB2-MB3-MB4	2
69396	MB1-MB3-MB4	1.8
69797	MB1-MB3-MB4	1.9
69594	MB1-MB3-MB4	1.9
69522	MB1-MB3-MB4	2

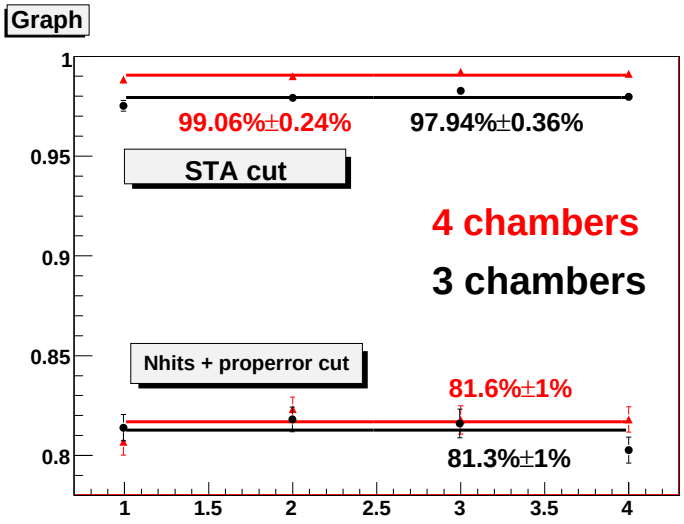

Figure 3.50: Proportion of events with at least one STA track for the runs listed in Tab. 3.3 in two cases: no requirement at all or some quality cuts imposed. In red runs with four active chambers, in black with three chambers. Data are fitted to a constant straight line.

4

Muon momentum measurement in CMS

In Chapter 3 it was shown that CMS aims to obtain numerous physics results through the analysis of physics events with muons in the final state. As explained there, the muon transverse momentum is determined in CMS by measuring the curvature of the particle inside the 3.8 T magnetic field created by the solenoid.

The path followed by the muon inside the solenoid is reconstructed by the inner tracker. The muon spectrometer located outside the solenoid and inserted in the iron return yoke of the magnetic field provides a measurement of the muon trajectory in that region.

Thus, two independent measurements of the muon momentum are available (19). Optimal resolution is reached when information from both subdetectors is combined to reconstruct the so-called global muon track (36), from which the muon momentum is derived. For muons with a transverse momentum $\lesssim$ 200 GeV, the resolution is, nevertheless, dominated by the resolution of the inner tracker measurement.

The muon momentum measurement can be affected by different factors (variation of the magnetic field with respect to the expected value, bad alignment,...). Important physical analyses (e.g. those based on the reconstructed dimuon invariant mass) depend on these deviations, imposing the necessity of knowing them as best as possible.

An accurate simulated MC description of the CMS detector is crucial for the analysis of many physics topics. Comparison between observed data distributions and predicted ones, in the framework of different physics models will allow to confirm or discard some

of them. To validate MC predictions, agreement data-MC, already for the very basic quantities, such as single muon distributions, should be optimal.

Extensive MC validation studies during the commissioning of the detector and in particular analyzing cosmic muon data allowed to reach a very good level of agreement data-MC, already before the startup of the LHC running.

In proton-proton collisions the channel pp $\rightarrow$ Z+$X \rightarrow \mu^+\mu^-+X$ allows the selection of very clean muon samples, once the dimuon invariant mass constraint is applied. This channel is thus an excellent candidate to improve and define the ultimate performance of CMS for muon momentum measurement, as well as to tune the MC description.

In this Chapter we explore the possibility to improve the MC muon momentum description determining additional effects, not accounted for in the MC, regarding muon p_T scale and resolution. The method proposed (hereafter called SIDRA, SImulation DRiven Analysis) is to fit the MC reconstructed Z mass distribution to the experimental one including "ad-hoc" functions that modify the MC reconstructed momentum. The parameters of these functions are the parameters to be fitted. The validity of the method is tested first on MC data. Afterwards it is applied on data collected during run 2010. Initial studies were driven for a $\sqrt{s} = 10$ TeV scenario (MC tests). However, data taken at LHC used bunches of protons at 7 TeV. Data analyzed at 7 TeV confirmed the predictions given by the tests at 10 TeV.

4.1 Muon momentum measurement

A charged particle moving in a magnetic field is subjected to a force $q\vec{v} \times \vec{B}$ where q is the charge of the particle, $\vec{v}$ is the velocity of the particle and $\vec{B}$ the magnetic field (38). In modern detectors, the momentum of charged particles is measured using high magnetic fields. The charged particles created in the collision are bent due to the effect of a magnetic field. In a solenoidal magnetic field, as the one CMS uses, the trajectory of these particles is circular in the transverse plane. The radius of the circular trajectory depends on the velocity of the particle

$$R = \frac{p_T}{0.3 \cdot q \cdot B} \tag{4.1}$$

with p_T being the transverse momentum (in GeV) of the particle, B the magnetic field (in Teslas), R the radious (in meters) and q the charge of the particle in proton charge units.

By measuring the radius of curvature of the circular trajectory we can measure the momentum of the charged particle in this plane (transverse momentum) knowing the strength of the magnetic field (B).

The accuracy of transverse momentum measurement of a charged particle inside a magnetic field is influenced by a number of different effects.

The intrinsic uncertainty is due to the finite track position resolution of the device. The size of this effect is constant in $1/p_T$,

$$\Delta(1/p_T) = \kappa$$

and thus, the relative resolution increases linearly with the momentum of the particle,

$$\frac{\Delta p_T}{p_T} = \kappa \cdot p_T$$

Other effects contribute to degrade the momentum measurement. Multiple scattering effects or magnetic field uncertainties induce an additional relative uncertainty that is essentially constant with the momentum,

$$\frac{\Delta p_T}{p_T} = \beta$$

Finally, if the particle traverses a large quantity of material, it may suffer strong energy losses, essentially independent of momentum. The effect on the uncertainty will be:

$$\frac{\Delta p_T}{p_T} = \frac{\gamma}{p_T}$$

In the most general case, all these effects will have an impact on the measurement, and thus the general expression for the uncertainty will be,

$$\frac{\Delta p_T}{p_T} = \kappa \cdot p_T \oplus \beta \oplus \frac{\gamma}{p_T}$$

The expected muon momentum resolution in CMS is presented in Fig. 4.1, both for the central part of the detector (left panel) and for the forward/backward regions (right panel). Predictions are shown for muons reconstructed only with the inner tracker detector, with the muon system and when both are combined. The resolution is $\lesssim 1\%$, in the whole η interval, for the low and medium momentum range ($p < 100$ GeV) (it is limited to $\sim 0.8\%$ by multiple scattering effects at low momentum), increasing to $\sim 10\%$ for high momentum ($p \simeq 1$ TeV).

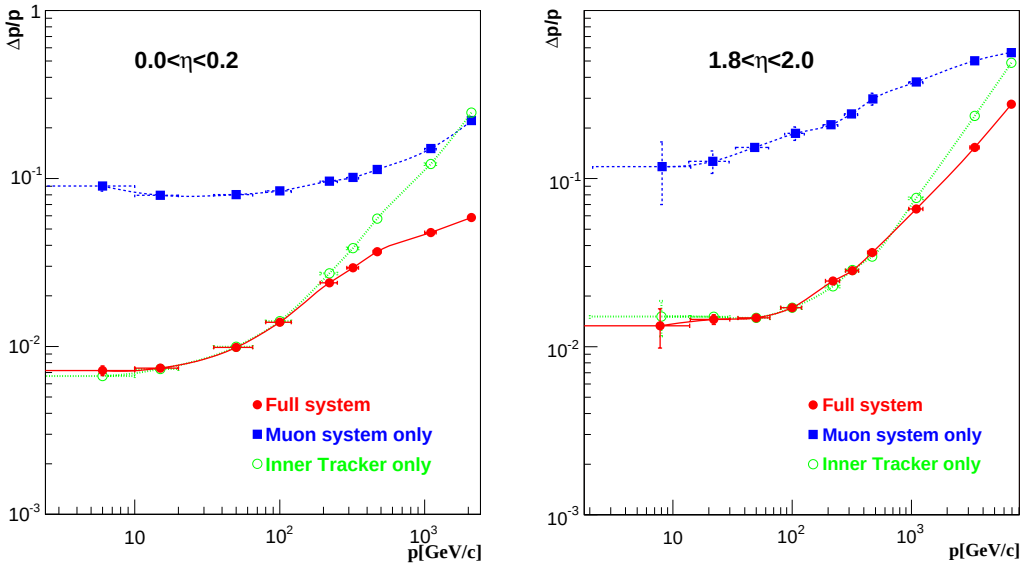

Figure 4.1: Expected muon momentum resolution in CMS. (Left) Barrel region. (Right) Endcap region.

The resolution for muon momentum measurement only in the muon system varies from 10 to $20-30\%$, depending on the momentum and pseudorapidity region.

Fig. 4.2 presents the reconstructed dimuon invariant mass distribution, obtained with a large pp $\rightarrow$ Z $+ X \rightarrow \mu^+\mu^- + X$ MC sample and it constitutes the expected experimental distribution once a Z-optimized selection process is applied on the simulation. This simulation is composed of a MC generator to emulate the collision process (hard interaction), the parton shower and hadronization, and a detector simulator. The expected performance of CMS as a muon spectrometer is reproduced with the GEANT4 package (39). The complete CMS geometry, including construction tolerances and the detailed description of the materials, together with the magnetic field map are included in the program.

Any inaccuracies in the simulation, such as alignment errors, discrepancies in the magnetic field value, incorrect evaluation of the material budget, etc. will give rise to a discrepancy between the simulated and the experimental distributions. Agreement data-MC can be retrieved if the reconstructed momentum in the simulation is further modified according to the expressions that describe the physical effects addressed before.

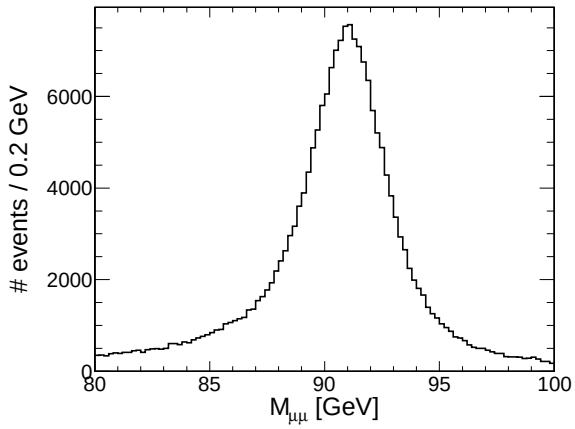

Figure 4.2: Reconstructed invariant mass distribution from a high statistics (1M event) simulated pp $\rightarrow$ Z + $X \rightarrow \mu^+\mu^- + X$ sample.

Muon momentum tuning will reflect in the simulated Z invariant mass distribution.

In this Chapter we have considered two possible effects:

- Scaling factors in the muon momentum. They induce a shift in the peak position of the reconstructed versus real momenta distribution.
- Extra resolution effects. They produce a widening of the reconstructed versus real momenta distribution.

Moreover, depending on the origin of these effects, they will impact on the measurement either of p_T or of $1/p_T$, thus leading to different parametrization in terms of the muon momentum:

1. Global scaling factor in the muon momentum, due for instance, to a global scaling factor in the magnetic field,

$$p_T\prime = (1 + \delta) \cdot p_T \tag{4.2}$$

2. An extra smearing term constant with momentum, as the one due to local differences in the magnetic field map or in the multiple scattering effect,

$$\frac{\Delta p_T}{p_T} = \sigma_{MS} \tag{4.3}$$

3. A constant shift in $1/p_T$,

$$\left(\frac{1}{p_T}\right)' = \frac{1}{p_T} + \delta_a$$

that translates into a p_T-scaling effect linearly increasing with the muon momentum,

$$p_T{}' = (1 + \delta_a \cdot p_T) \cdot p_T \tag{4.4}$$

4. An extra resolution effect, following the same parametrization as an intrinsic resolution term, i.e. constant in $1/p_T$,

$$\Delta(1/p_T) = \sigma_{intrinsic}$$

It translates into a smearing factor linearly depending on the momentum,

$$\frac{\Delta p_T}{p_T} = \sigma_{intrinsic} \cdot p_T \tag{4.5}$$

Studies performed with low mass resonances samples have shown that the effects constant in p_T are negligible (40, 41). For this reason, only effects linear in p_T will be considered hereafter (they will be referred as $\delta(1/p_T)$ and $\sigma(1/p_T)$ in this Chapter).

Figs. 4.3 and 4.4 illustrate how the reconstructed dimuon-invariant mass distribution is affected by these two effects presented above. In the case of Fig. 4.3 the reconstructed dimuon invariant mass distribution is shown, where the transverse momentum of every muon is scaled with two different linearly dependent scaling factors ($\delta(1/p_T) = 0.03$ TeV^{-1} and 0.3 TeV^{-1}). Fig. 4.4 illustrates the case where the reconstructed p_T of every muon is smeared according to an intrinsic resolution-like term ($\sigma(1/p_T) = 0.6$ TeV^{-1}).

The main effect of a scaling factor is a displacement of the peak position with respect to the original one. In addition to this shift, the result of a p_T-dependent scale factor is that the distribution is somewhat widened in an asymmetric way, more visible in the case of $\delta(1/p_T) = 0.3$ TeV^{-1}.

The net effect of the additional smearing term is a visible widening of the reconstructed dimuon mass distribution, remaining the peak almost in the same position. In general, all these effects may affect the muon momentum measurement at the same time, being the peak interval (88 GeV $< m_{\mu\mu} <$ 94 GeV) the most sensitive region to them.

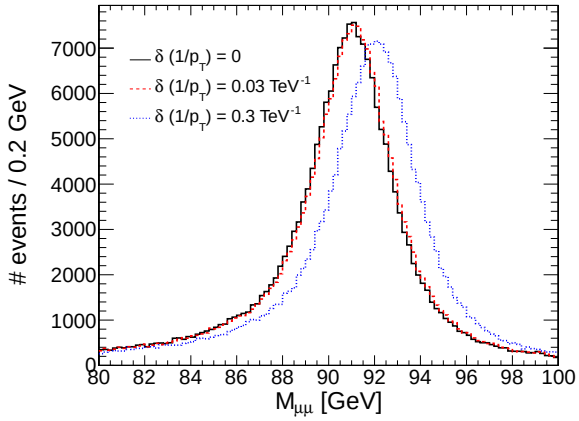

Figure 4.3: Reconstructed invariant mass distribution from a high statistics (1M events) simulated pp $\rightarrow$ Z + X $\rightarrow$ $\mu^+\mu^-$ + X sample. The p_T of every muon is scaled with a scaling factor δ linearly dependent on the muon p_T. Two values of δ are shown: $\delta(1/p_T) = 0.03$ TeV^{-1} (red, dashed histogram) and $\delta(1/p_T) = 0.3$ TeV^{-1} (blue, dotted histogram). The original distribution ($\delta(1/p_T) = 0$) is shown as a solid black line.

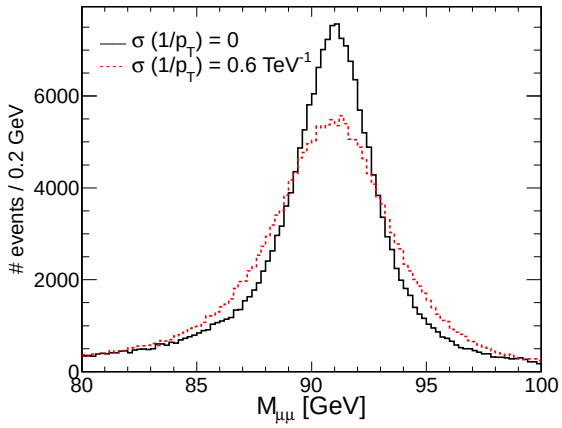

Figure 4.4: Reconstructed invariant mass distribution from a high statistics (1M events) simulated pp $\rightarrow$ Z + X $\rightarrow$ $\mu^+\mu^-$ + X sample. The p_T of every muon is smeared with an additional Gaussian function of width $\sigma(1/p_T) = 0.6$ TeV^{-1} (red, dashed histogram). The original distribution ($\sigma(1/p_T) = 0$) is shown as a solid black line.

4.2 Method

At generator level the Z-dimuon invariant mass distribution is well described by the different MC generators and hardly any differences are appreciated in the Z peak neither between LO and NLO predictions, nor due to PDF effects. Several MC studies have shown the robustness of the simulated Z mass distribution, showing very little variations due to the different theoretical assumptions to generate it. Therefore, experimental effects will be the only ones significantly distorting the Z mass shape. Among them, the most relevant ones will be muon scale and resolution and they will have a direct impact in the reconstructed Z-mass distribution.

The method proposed is to fit the MC reconstructed Z mass distribution to the experimental one: the muon description in the MC is allowed to vary according to certain functions and the following binned likelihood function, arising from a Poissonian distribution, is minimized:

$$-\log L = \sum_{i=1}^{i=nbins} [N_i^{\mathrm{MC}}(\delta(1/p_T), \sigma(1/p_T)) - n_i^{\mathrm{data}} \times \log N_i^{\mathrm{MC}}(\delta(1/p_T), \sigma(1/p_T))] \quad (4.6)$$

where the index i runs over all bins in the invariant mass distribution between 70 GeV and 110 GeV, in 1 GeV bins. The variables n_i^{data} and $N_i^{\mathrm{MC}}(\delta(1/p_T), \sigma(1/p_T))$ are the number of entries in the i mass bin of the experimental and MC distributions, respectively.

The p_T of the muon is distorted using the previous expressions. The MC Z-mass distribution is consequently modified and for every $m_{\mu\mu}$ bin, N^{MC} is a function of the parameters modifying the reconstructed momentum, $N^{\mathrm{MC}} = N^{\mathrm{MC}}(\delta(1/p_T), \sigma(1/p_T))$. $\delta(1/p_T)$ and $\sigma(1/p_T)$ are the parameters of the minimization.

Fig. 4.5 illustrates the method. A small subset of the MC Z-sample is taken to play the role of the data ("fake data" sample), in this case of a 1 pb^{-1} equivalent luminosity. Fig. 4.5 (left) shows the Z reconstructed distribution from this "fake data" sample, where the p_T of the muons is distorted simulating a possible experimental case. In this example a smearing effect has been simulated. The disagreement between the "fake data" and MC can be clearly noticed.

The p_T of the muons in the MC sample is now described in terms of the functional forms defined before, and expression (4.6) is minimized in order to recover the distortion introduced.

Fig. 4.5 (right) shows the same "fake data" distribution and the MC distribution resulting from the minimization. The agreement between fitted MC and "fake data" is significantly improved. The extra smearing introduced in the fake data sample ($\sigma(1/p_T) = 1.5$ TeV^{-1}) is recovered by the fit, as $\sigma^{\text{fitted}}(1/p_T) = 1.53$ TeV^{-1}.

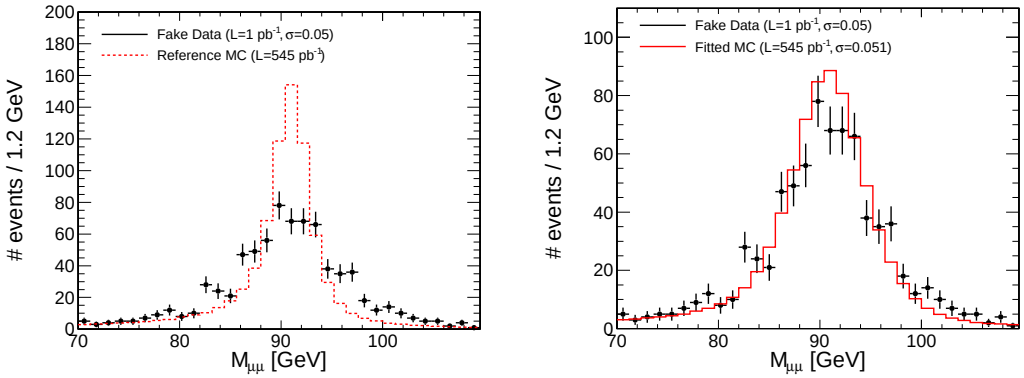

Figure 4.5: "Fake data" Z invariant mass distribution (black dots) and MC Z invariant mass distribution (red histogram). (Left) The MC prediction is produced with ideal detector conditions. (Right) The MC prediction is fitted to the "fake data" sample including terms distorting the p_T of the muons.

4.3 Validation of the method with MC samples

The method was established with a high statistics MC sample of pp $\rightarrow$ Z+$X \rightarrow \mu^+\mu^- + X$ events at $\sqrt{s} = 10$ TeV, generated with Pythia (42, 43). From this sample, smaller sub-samples were selected to play the role of fake data sets. The data process, generator, number of events generated, cross section and equivalent integrated luminosity of the process are summarized in Table 4.1.

Z-boson candidates are selected from those events passing the single muon trigger path. Events are required to have two opposite charge, high momentum ($p_T > 20$ GeV) muons in the pseudorapidity range $|\eta| < 2$.

Certain quality criteria are imposed to reconstructed muons (46): the muon is required to have been reconstructed with two different reconstruction algorithms. The χ^2/ndf of the global muon track fit should be less than 10 and the number of hits used to fit the track in the inner part of the detector should be greater than 10. The impact

Table 4.1: MC data sample used in this study. Cross section quoted include the Branching Ratio of Z→ $\mu^+\mu^-$. The Z cross section is computed at NLO with MCFM (44) program and MSTW08 (45) NLO Central PDF.

Process	Generator	**# events**	σ (pb)	$\mathcal{L}_{\text{int}}$ (pb^{-1})
$\text{pp} \to \text{Z} + \text{X} \to \mu^+\mu^- + \text{X}$	PYTHIA	1.3 M	2403	545

parameter in the transverse plane is required to be below 0.2 cm. In addition, muons must be isolated. A relative isolation definition is used; it is defined as the sum of the p_T of the tracks, in a cone of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$, centered in the muon track, and normalized to the muon p_T. A muon is considered to be isolated if the isolation variable is smaller than 0.1. These requirements are set in order to get a clean Z boson sample. More details about this set of cuts can be seen in Chapter 6. In Table 4.2 the number of events remaining in the MC samples, after the different requirements is shown. The last column of the table indicates the expected number of selected events for an integrated luminosity of 1 pb^{-1}.

Table 4.2: Number of events after the different steps in the selection process. The **Acceptance** column indicates the number of events with both muons in the acceptance region $|\eta| < 2$, and the **Two sel. muons** column refers to the number of events containing two muons of opposite charge and fulfilling the quality and selection requirements (high p_T and isolation for $\text{Z} \to \mu^+\mu^-$ selection). Last column gives the expected number of selected events for an integrated luminosity of 1 pb^{-1} at 10 TeV.

Process	**# events**	**Trigger**	**Acceptance**	**2 sel. μs**	**#ev./pb^{-1}**
$\text{pp} \to \text{Z} + \text{X} \to \mu\mu + \text{X}$	1309714	622940	351795	227890	418

With the presented selection requirements, the residual backgrounds are due to other electroweak processes (such as $W \to \mu\nu$ or $Z \to \tau\tau$) and $t\bar{t}$. However, the total background contribution estimated using MC samples is negligible (amounts to less than 1% of the final sample) and can be ignored.

4.4 Results on MC using fake data sets

We have established the performance of the method, from a quantitative point of view, with the MC sample. Muons in the "fake data" sample have been distorted according to the effects presented before (with different, known, input parameters) and their dimuon invariant-mass distribution computed. The reconstructed dimuon distribution from the remaining MC sample is fitted to this "experimental" one and the output parameters are compared to the input ones. In this part, two luminosity scenarios are explored. This method was initially developed for the beginning of the data taking, in a low statistics scenario. A sample of $\mathcal{L}_{\rm int} = 1$ pb^{-1} serves to emulate a low statistics scenario, for the very beginning of the data taking. However, most of the analyses will benefit from a bigger amount of data. A second scenario (with a luminosity of $\mathcal{L}_{\rm int} = 10$ pb^{-1}) is explored to check the capabilities of the method in these conditions.

The initial multidimensional problem reduces to a one-dimensional one if only one of the effects distorting the sample is considered. Fig. 4.6 shows the $-$(log-likelihood) curves for two particular cases of the one-dimensional problems: a p_T-dependent scale factor term with $\delta(1/p_T) = 1. \times 10^{-2}$ TeV^{-1} and an additional intrinsic-like term with an input $\sigma(1/p_T) = 2. \times 10^{-2}$ TeV^{-1}.

In all cases the $-$(log-likelihood) functions are parabolic around the minimum. For the smearing effect, it exhibits a double parabola structure symmetrically located with respect to a null distortion parameter[1].

Both minima will get closer when dealing with distortion values close to zero, making a rather broad minimum in the limit of a non-distorted sample ($\sigma \sim 0$).

To test the method for a 1 pb^{-1} luminosity scenario, and given the simulated MC statistics (equivalent to 545 pb^{-1}), 545 independent tests with 1 pb^{-1} luminosity subsets behaving as "fake data" can be made, taking in each case the remaining of the Z sample as the MC sample to be fitted. To test the method for a 10 pb^{-1} luminosity scenario, the available MC statistics only affords 54 independent tests.

[1] Due to the nature of the fit, positive and negative values of the sigma parameter are equally probable.

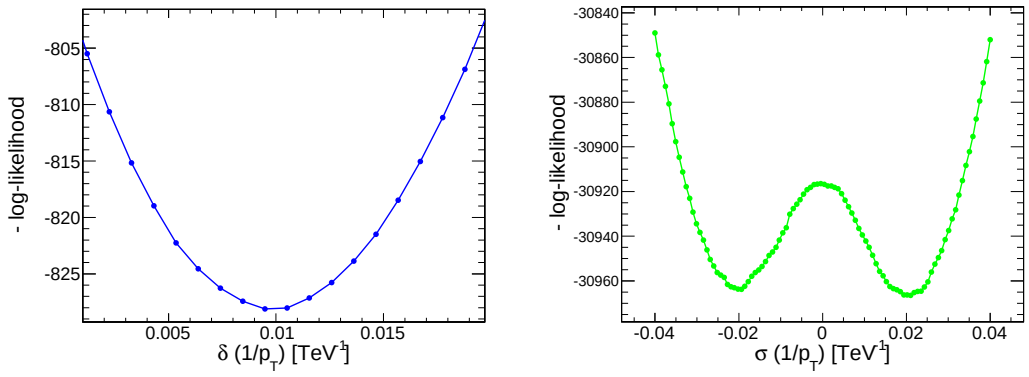

Figure 4.6: $-$(Log-likelihood) curves for: (left) a scale factor linearly dependent with the momentum $\delta(1/p_T) = 1. \times 10^{-2}$ TeV^{-1} and (right) an additional intrinsic-like term with $\sigma(1/p_T) = 2. \times 10^{-2}$ TeV^{-1}.

4.4.1 Constant shift in $1/p_T$

The p_T of the muons is modified with a constant shift in the inverse of the momentum. This effect is equivalent to a linearly p_T-dependent scale factor:

$$p_T\prime = [1 + \delta(1/p_T) \cdot p_T] \cdot p_T \tag{4.7}$$

This modification is applied to muons in "fake data" samples of 1 and 10 pb^{-1} integrated luminosity, and the full MC is fitted to them, trying to retrieve the input distortion parameters.

The fitted scale factors when dealing with 1 pb^{-1} luminosity data sets are plotted in Fig. 4.7 (left) and Fig. 4.8 (left), once they are distorted with scale factors $\delta(1/p_T) = 0.3$ TeV^{-1} and $\delta(1/p_T) = 0.03$ TeV^{-1}, respectively.

The case of no distortion ($\delta(1/p_T) = 0$) is presented in Fig. 4.9. The right plots in the same Figs. show the distribution of the fitted parameters. The mean values are in agreement with the input value in all three cases.

The width of these distributions are of the order of 0.041 TeV^{-1}. They are in agreement with the expected error values from the $-$(log-likelihood) curves, like the one plotted in Fig. 4.10, shown as an example.

Table 4.3 summarizes the recovered values (mean and width of the distributions of fitted parameters) for three different input values of the scale factor, and for the two luminosity scenarios considered. We can conclude that an overall linearly p_T-dependent

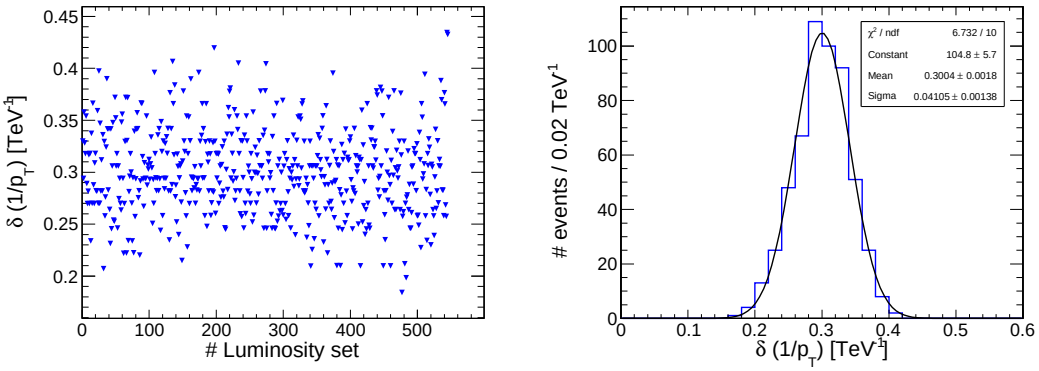

Figure 4.7: (Left) Scale factor recovered for each 1 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0.3$ TeV^{-1}.

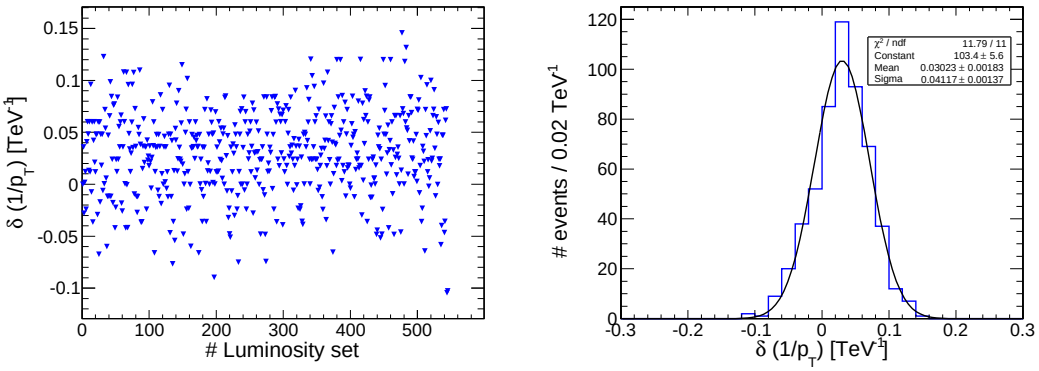

Figure 4.8: (Left) Scale factor recovered for each 1 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0.03$ TeV^{-1}.

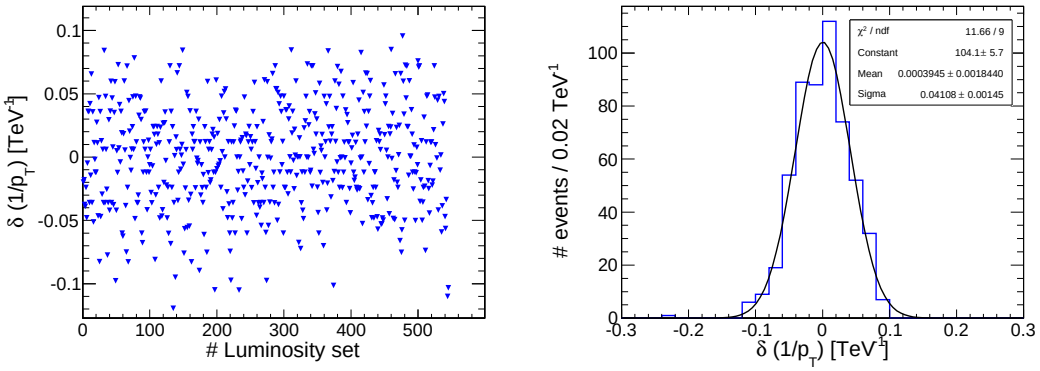

Figure 4.9: (Left) Scale factor recovered for each 1 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0$.

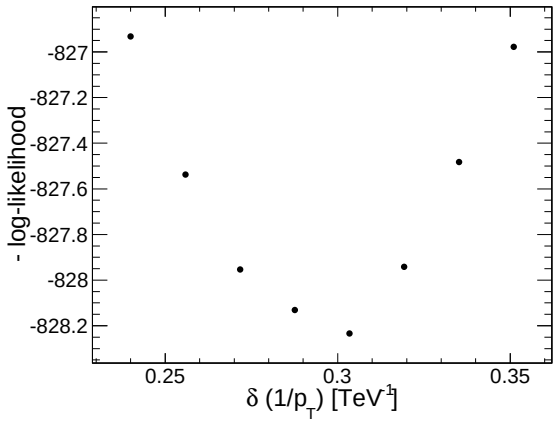

Figure 4.10: $-$(Log-likelihood) shape for the p_T-dependent scale factor term for a particular 1 pb^{-1} "fake data" luminosity set when the "fake data" is distorted with an input scale factor $\delta(1/p_T) = 0.3$ TeV^{-1}. The error interval defined by the increase of the $-$log-likelihood in 0.5 from the minimum units is $[0.264, 0.33]$ TeV^{-1}, from which an average error of $\Delta\delta(1/p_T) = 0.033$ TeV^{-1} is derived.

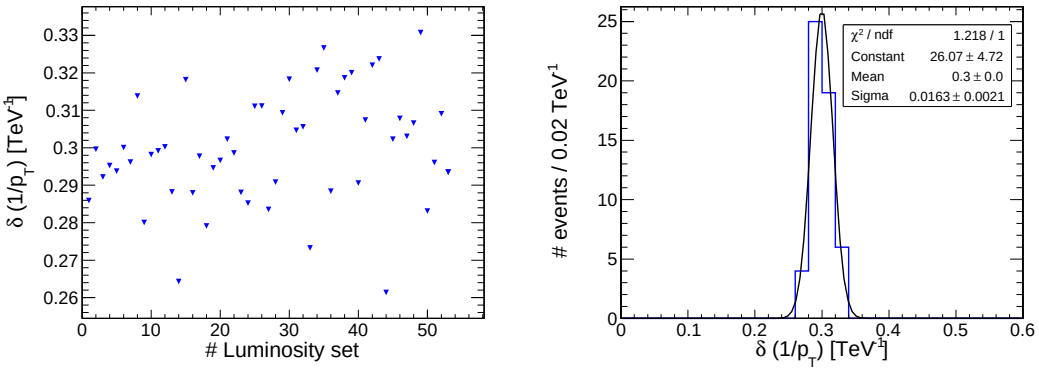

Figure 4.11: (Left) Scale factor recovered for each 10 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0.3$ TeV^{-1}.

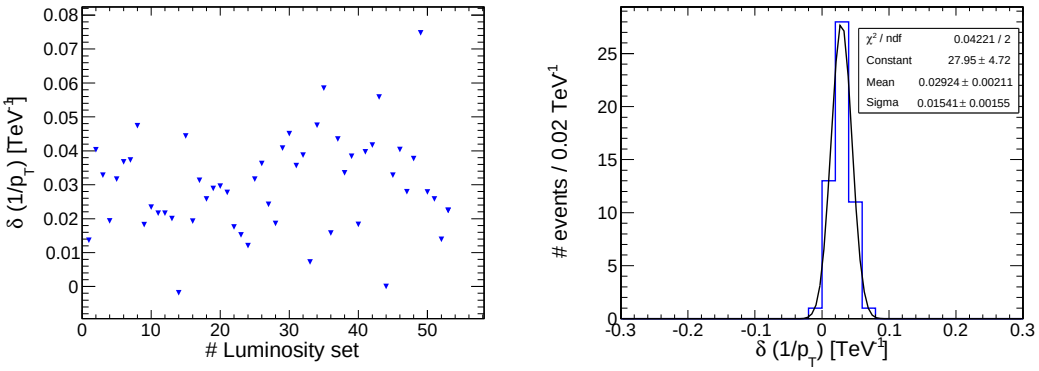

Figure 4.12: (Left) Scale factor recovered for each 10 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0.03$ TeV^{-1}.

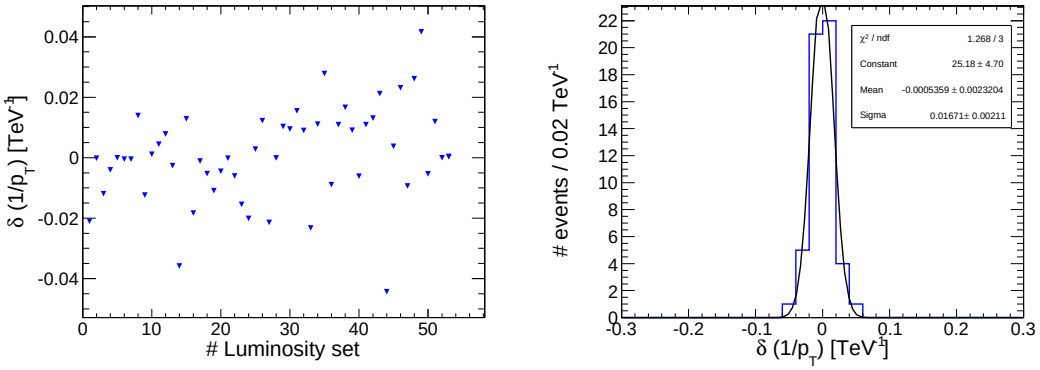

Figure 4.13: (Left) Scale factor recovered for each 10 pb^{-1} "fake data" luminosity set. (Right) Histogram of fitted values. The input value is $\delta(1/p_T) = 0$.

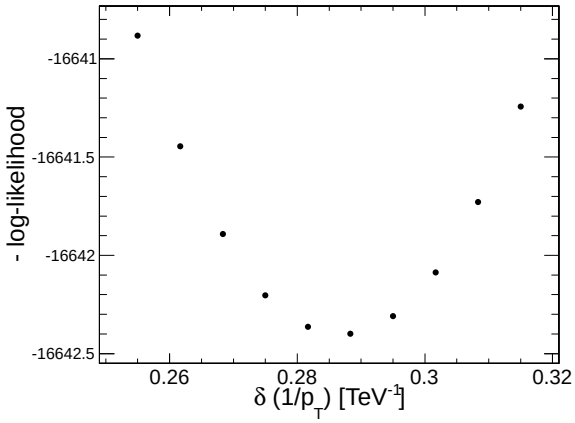

Figure 4.14: $-$(Log-likelihood) shape for the p_T-dependent scale factor term for a particular 10 pb^{-1} "fake data" luminosity set for an input scale factor $\delta(1/p_T) = 0.3$ TeV^{-1}. The error interval defined by the increase of the $-$log-likelihood in 0.5 units from the minimum is $[0.267, 0.303]$ TeV^{-1}, from which an average error of $\Delta\delta(1/p_T) = 0.018$ TeV^{-1} is derived.

scale factor can be determined, if it is larger than ~ 0.041 TeV^{-1}, even with very low collected luminosity ($\mathcal{L}_{\rm int} \sim 1$ pb^{-1}).

Table 4.3: Mean and width of the distributions of fitted scale factor parameter (all magnitudes are in TeV^{-1}).

	$\mathcal{L}_{\rm int} = 1$ pb^{-1}		$\mathcal{L}_{\rm int} = 10$ pb^{-1}	
$\delta_{\rm input}(1/p_T)$	$< \delta_{\rm fitted}(1/p_T) >$	$\sigma(\delta_{\rm fitted}(1/p_T))$	$< \delta_{\rm fitted}(1/p_T) >$	$\sigma(\delta_{\rm fitted}(1/p_T))$
0.3	0.3	0.041	0.3	0.016
0.03	0.03	0.041	0.03	0.015
0.	3.94×10^{-4}	0.041	-5.36×10^{-4}	0.017

4.4.2 Intrinsic resolution-like term

Muon p_T in the "fake-data" sample is now distorted with a term of the form:

$$\Delta p_T/p_T = \sigma(1/p_T) \cdot p_T$$

The $\sigma(1/p_T)$ parameters obtained when the MC is fitted to the different 1 pb^{-1} "fake-data" luminosity sets are presented in Fig. 4.15, Fig. 4.16 and Fig. 4.17. Input values are $\sigma(1/p_T) = 1.5$ TeV^{-1}, $\sigma(1/p_T) = 0.6$ TeV^{-1} and no distortion. The histogram of the fitted values is also shown in the same Figs. The mean values agree well with the input ones, with a precision of ~ 0.1 TeV^{-1}.

In all cases there is an agreement between the expected precision and the behavior of the likelihood function (Fig. 4.18 shows a particular example). Here again, the fact that the two parabolas of the $-$(log-likelihood curve) approach to each other (see Fig. 4.18) when trying to fit in the non distortion case leads to a slightly higher uncertainty (~ 0.15 TeV^{-1}).

The same exercise is repeated with 10 pb^{-1} "fake-data" luminosity samples. The results are presented in Fig. 4.19, Fig. 4.20 and Fig. 4.21. Mean values of the fitted parameters distributions agree well with the input values. Concerning precision, a factor of 2 to 3 of improvement is observed, depending on the case. Now, there is also an agreement between the width of the histogram and the behavior of the $-$(log-likelihood) curve (like the one shown in Fig. 4.22).

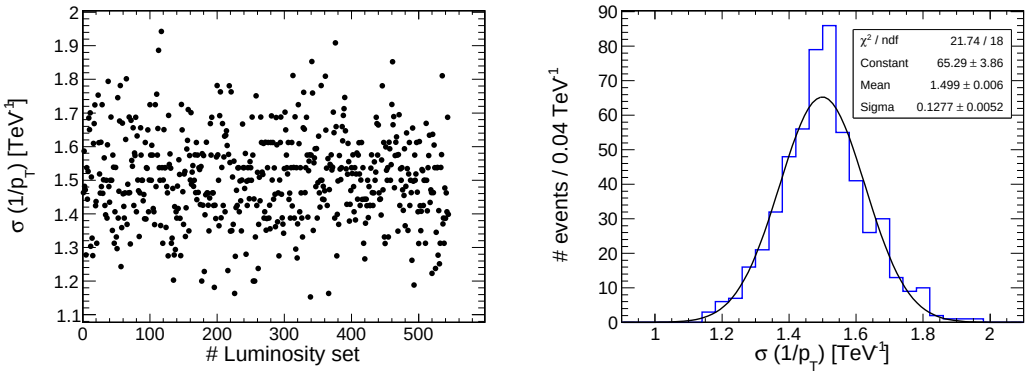

Figure 4.15: (Left) Additional intrinsic-like term recovered for each 1 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 1.5$ TeV^{-1}.

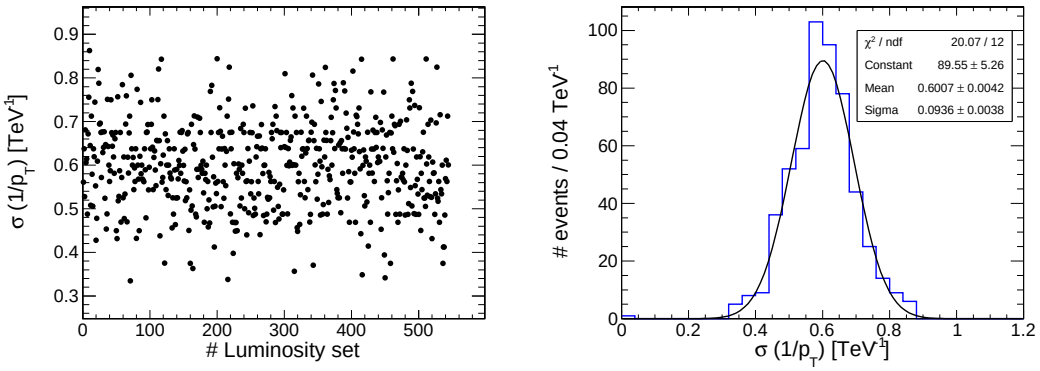

Figure 4.16: (Left) Additional intrinsic-like term recovered for each 1 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 0.6$ TeV^{-1}.

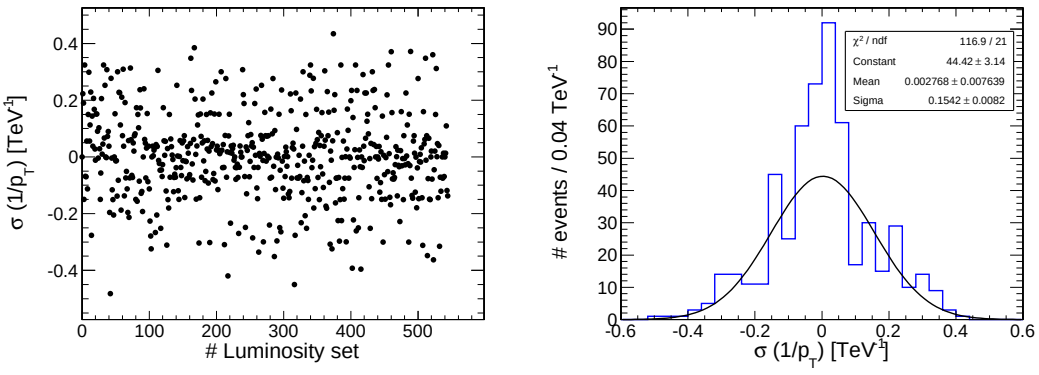

Figure 4.17: (Left) Additional intrinsic-like term recovered for each 1 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 0$.

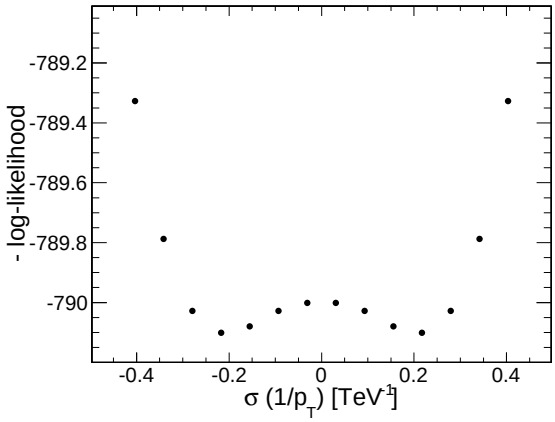

Figure 4.18: $-$(Log-likelihood) shape for the intrinsic-like smearing term for a particular 1 pb^{-1} luminosity set for a $\sigma(1/p_T) = 0$ input. As the $-$log-likelihood function has two minima, the error in the parameter is estimated checking the parameter value where the $-$log-likelihood function increases in 0.5 units from the minimum, evaluated to the right of the minimum (at $\sigma(1/p_T) \sim 0.21$ TeV^{-1}). This increase is reached at $\sigma(1/p_T) = 0.36$ TeV^{-1}, and an average error of $\Delta\sigma(1/p_T) = 0.15$ TeV^{-1} is derived.

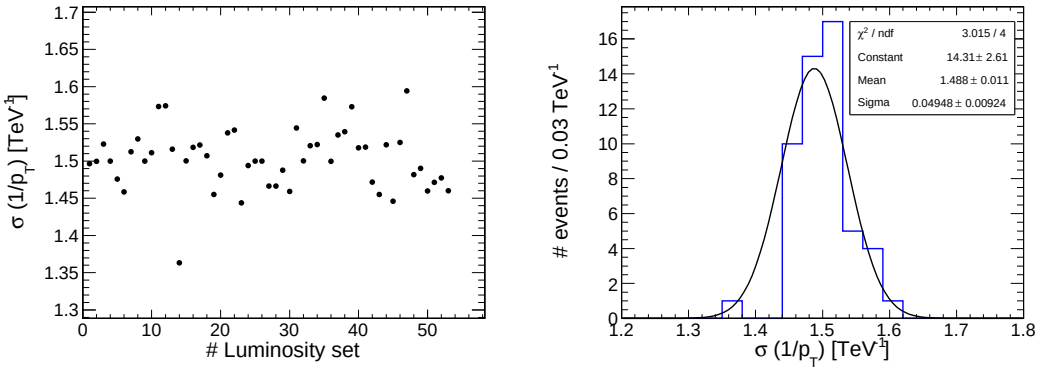

Figure 4.19: (Left) Additional intrinsic-like term recovered for each 10 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 1.5$ TeV^{-1}.

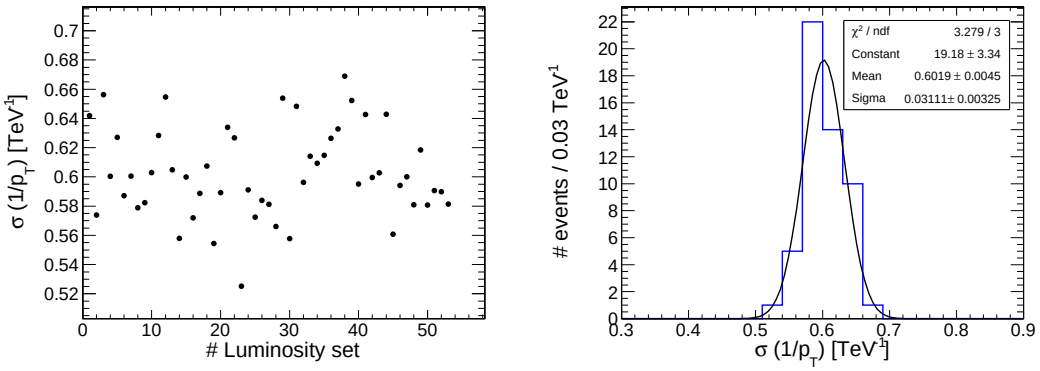

Figure 4.20: (Left) Additional intrinsic-like term recovered for each 10 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 0.6$ TeV^{-1}.

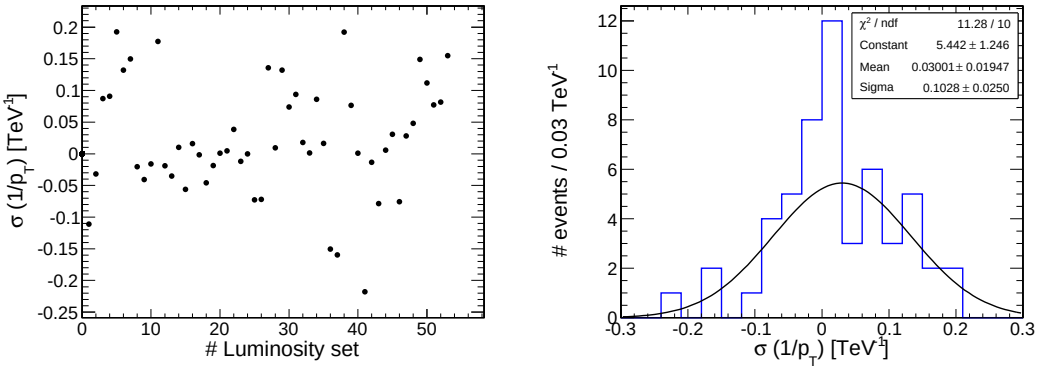

Figure 4.21: (Left) Additional intrinsic-like term recovered for each 10 pb^{-1} luminosity set. (Right) Histogram of fitted values. The input value is $\sigma(1/p_T) = 0$.

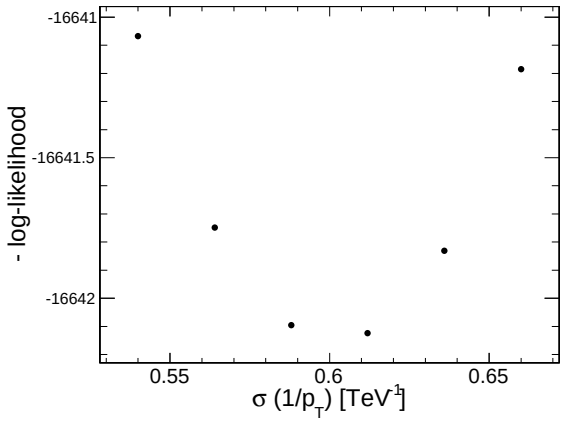

Figure 4.22: $-$(Log-likelihood) shape for the additional intrinsic-like term for a particular 10 pb^{-1} luminosity set with a $\sigma(1/p_T) = 0.6$ TeV^{-1} input. The error interval defined by the increase of the $-$log-likelihood in 0.5 units is $[0.571, 0.642]$ TeV^{-1}, from which an average error of $\Delta\sigma(1/p_T) = 0.035$ TeV^{-1} is derived.

Table 4.4 summarizes the recovered values (mean and width of the distributions of fitted parameters) for three different input values of an additional smearing term, with the dependence on the muon momentum like an intrinsic resolution term, and for the two luminosity scenarios considered. We can conclude that an overall intrinsic resolution-like term larger than ~ 0.12 TeV^{-1}, can be well recovered, even with very low collected luminosity ($\mathcal{L}_{\rm int} \sim 1$ pb^{-1}). It means that an additional intrinsic resolution contribution, larger than 15% TeV^{-1}, can be determined.

Table 4.4: Mean and width of the distributions of fitted parameter $\sigma(1/p_T)$ (all magnitudes are in TeV^{-1}.)

	$\mathcal{L}_{\rm int} = 1$ pb^{-1}		$\mathcal{L}_{\rm int} = 10$ pb^{-1}	
$\sigma_{\rm input}(1/p_T)$	$< \sigma_{\rm fitted} > (1/p_T)$	$\sigma(\sigma_{\rm fitted})(1/p_T)$	$< \sigma_{\rm fitted} > (1/p_T)$	$\sigma(\sigma_{\rm fitted})(1/p_T)$
1.5	1.5	0.129	1.49	0.049
0.6	0.6	0.093	0.602	0.031
0.0	27.7×10^{-3}	0.153	0.03	0.1

4.5 Results on collision data

Once the performance of the method was established in MC, it was applied to data already with the very first data collected. The final data sample used corresponds to 36 pb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV (all the data recorded during 2010). After the full selection process 12252 Z bosons are selected.

The MC used in this section corresponds to a sample generated with POWHEG (42). When applying the method to real data, the selection slightly differs from the one applied to the validity tests with MC to comply with the standard selection procedure set in the context of electroweak studies in CMS. In this case, no trigger is required, the pseudorapidity region considered is set to 2.1 and the isolation variable threshold is set to 0.15. Several tests confirmed the independence of the results with respect to these variations.

4.5.1 Misalignment scenarios on MC samples

Simulated events are reconstructed with the same programs used for real data. Specialized studies carried out in parallel with the physics studies, to assess the performance of the detector, provide a set of parameters that define the working conditions of the detector. These parameters are stored in a database and are incorporated into the reconstruction programs. A new reconstruction of both, real data and simulated events may be needed in case new effects are discovered or a new set of running conditions becomes available.

For muons coming from a Z decay (medium p_T range, where the resolution from the tracker dominates) the residual misalignments of the different elements that compose the tracker are specially relevant. These misalignment conditions are evaluated by software and included in the database for general use in the track reconstruction. The method consists basically of determining with a selected collection of tracks a set of new positions of the tracker modules such that it minimizes the residuals of the tracker hits with respect to the reconstructed track.

Even if the minimization converges and the pattern recognition performs correctly, possible remaining misalignment could still be present, biasing the measurements of the track parameters. In particular, those effects that keep invariant the χ^2 minimized by the alignment algorithm will not be detected by the procedure. These transformations of the geometry are the so called "weak modes" (47).

The simplest example of a weak mode is a global translation or rotation of the whole tracker. Any of these transformations will keep the χ^2 of the track invariant, not being spotted by the alignment algorithm. This trivial example is easily removed by other means, but several non-trivial transformations affecting the geometry of the tracker, could still be present after the track based alignment. Systematic biases in a physics measurement could arise as a consequence: a systematic rotation of the layers of the tracker would give rise to a charge-dependent momentum asymmetry in the muon tracks. In a similar way, the radial expansion or compression or a Z-scale distortion would change the distance scale measured, and, as a consequence, would distort any lifetime measurements.

For the lowest-order modes, 9 systematic misalignments for a cylindrical geometry are identified. They can be classified by the 3 different coordinates in which the shift

Table 4.5: Classification of weak modes by the way the tracker geometry is modified.

	ΔR	$\Delta\phi$	ΔZ
R	Radial expansion	Curl	Telescope
ϕ	Elliptical	Clamshell	Skew
Z	Bowing	Twist	Z expansion

can take place (ΔR, $\Delta\phi$ and ΔZ) and the dependence of this shift (in R, ϕ and Z) as can be seen in Tab. 4.5 and Fig. 4.23.

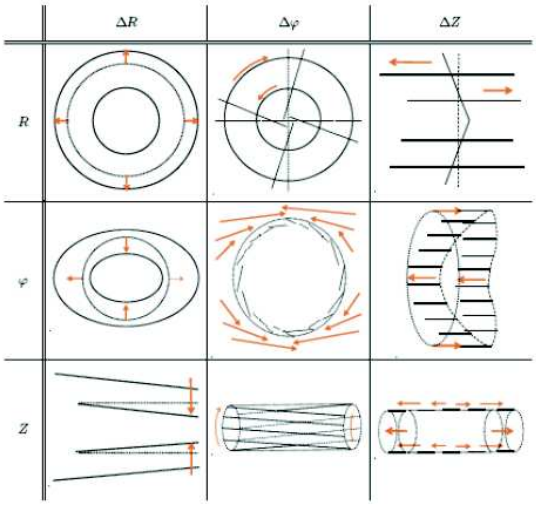

Figure 4.23: Weak modes - Graphical representation of the weak modes and their classification by the way the tracker geometry is modified.

In Fig. 4.24 the reconstructed momentum minus the generated momentum is plotted for the nine studied weak modes. With them, the global effect included in the startup geometry (last row, right plot, which is the one applied to MC) and a reduced offset in the transverse curvature $\kappa = q/p_T$ of 0.5 TeV^{-1} are shown. In this Fig. the misalignments are applied over a sample of simulated W bosons, which are selected using an optimized selection process. Several conclusions can be drawn from these plots: none of these modes shifts the muon momentum distribution, except the curl mode (this mode can be removed in most of the cases); all these modes make the muon momentum distribution wider; the reduced mode induces a sinusoidal behavior in ϕ.

In order to make the MC resemble the data, a misalignment scenario is simulated.

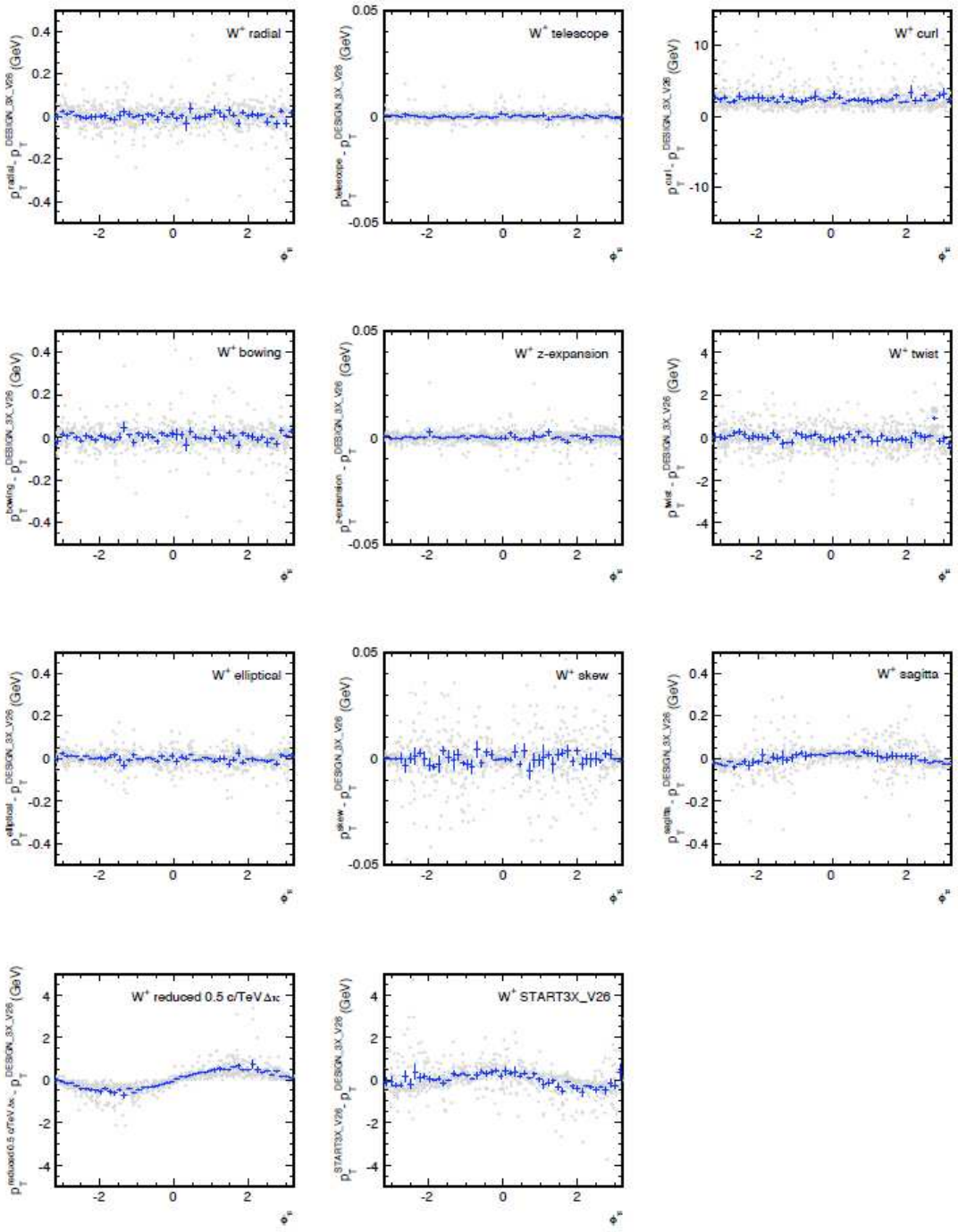

Figure 4.24: Impact of weak modes on muon p_T - Impact of the 9 studied weak modes on the muon transverse momentum of a simulated sample of W^+ bosons decaying into a muon and a neutrino (upper three rows). The lower two plots illustrate two additional geometry scenarios: the effect due to a reduced offset in the transverse curvature $\kappa = q/p_T$ of 0.5 TeV^{-1}, and the actual geometry used for the simulation and reconstruction of the data sample.

Due to their properties, previously described, the weak modes induced by this scenario are not fully removed after the alignment procedure. Among the possible modes that could still be present in the final simulated sample, the reduced one stands out. It produces a shift with sinusoidal shape in ϕ. In this mode, shifts in one hemisphere of CMS would be positive, and negative in the opposite hemisphere. Furthermore, they are opposite in sign for both charges. Since the muon distribution must be flat in ϕ, the net shift effect of this mode is null. The peak of the invariant mass distribution remains in the same region. However, a widening of this distribution can be observed when this mode is present, inducing a similar effect as the resolution term does.

In Fig. 4.25 the invariant mass distribution for the MC Z $\rightarrow \mu^+\mu^-$ sample versus the negative muon ϕ can be seen. A ϕ-dependent peak position is observed. This ϕ dependent shift makes the reconstructed dimuon mass peak wider.

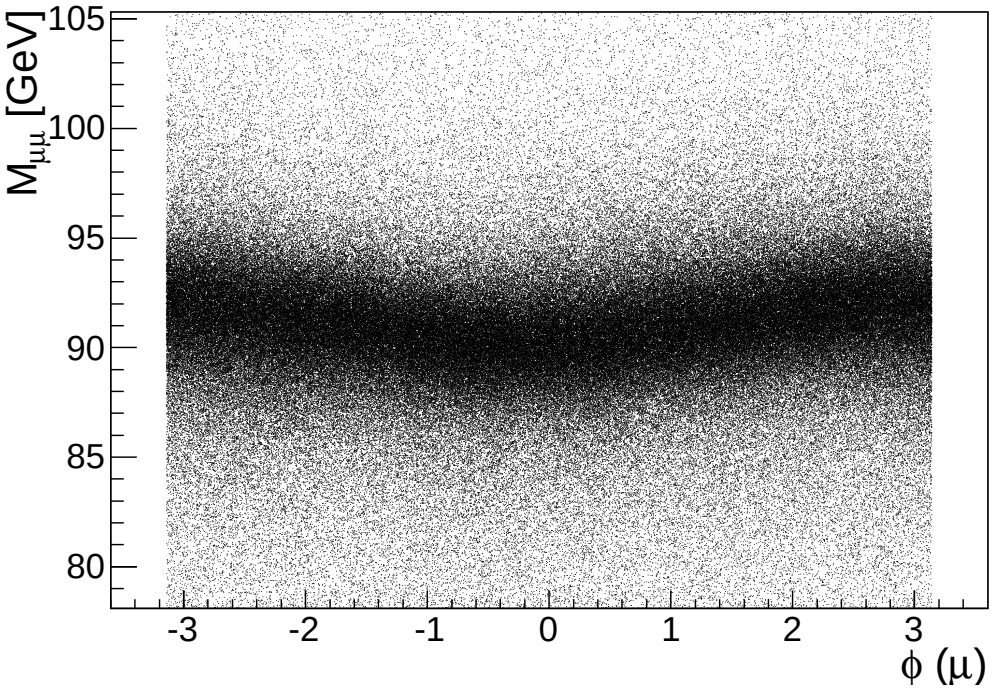

Figure 4.25: Z invariant mass distribution versus negative muon ϕ - A ϕ dependent shift is clearly observed in the Z invariant mass distribution. The reduced weak mode shown before is present in the simulated tracker geometry.

We have then followed a two-step procedure: First, we have applied the method over

the MC sample (using the generator information as the reference and the reconstructed momentum as "fake data") to remove these weak modes already present in the simulated samples. Afterwards, we have repeated the process but now comparing data with the already corrected MC. In the first case, we use a sinusoidal form to parametrize the shift in ϕ. After the fit we get:

$$\delta(1/p_T) = 0.00107 - 0.278 \cdot q \cdot \sin(\phi - 1.517)\ [\mathrm{TeV}^{-1}]$$

In Fig. 4.26 the result of this fit is plotted. In dots the Z boson peak using data is shown. In green the reference MC and in red the corrected MC are shown. After correcting the weak mode, the MC distribution is narrower than before, as expected following the previous reasoning. In this situation we can apply the method over the data.

As shown in this subsection, a parametrization of the scale factor in ϕ becomes necessary to remove the weak modes. In addition, and to reach a better agreement data-MC and a better description of the effects, a dependence of the scale factor in η is introduced. The resolution term is allowed to vary as a function of η.

4.5.2 Scale factor and resolution parametrization in η and ϕ

In this subsection we will derive a full parametrization of the shifts and resolution terms as a function of η and ϕ. We will assume that the resolution term does not depend on ϕ.

4.5.2.1 Scale factor dependence on ϕ and η

The scale factor will be studied first independently as a function of ϕ and η as a prior step to the full parametrization in the two variables. The full phase space is divided into several regions in ϕ and η and a single parameter, describing the scale factor will be determined for each of them, and for each muon charge. Six regions in the azimuthal angle are defined covering each of them 60 degrees in ϕ. In η, four regions are identified: barrel ($|\eta| < 0.9$), overlap ($0.9 < |\eta| < 1.2$), near endcap ($1.2 < |\eta| < 2.1$) and far endcap ($|\eta| > 2.1$) and also distinguishing between positive and negative η. The total number of parameters in the fit are 12 (for ϕ) plus 16 (for η).

In Fig. 4.27 the fitted scale factor in every ϕ region is plotted. From this plot we can conclude that: the shifts in ϕ follow a sinusoidal shape for both positive and negative

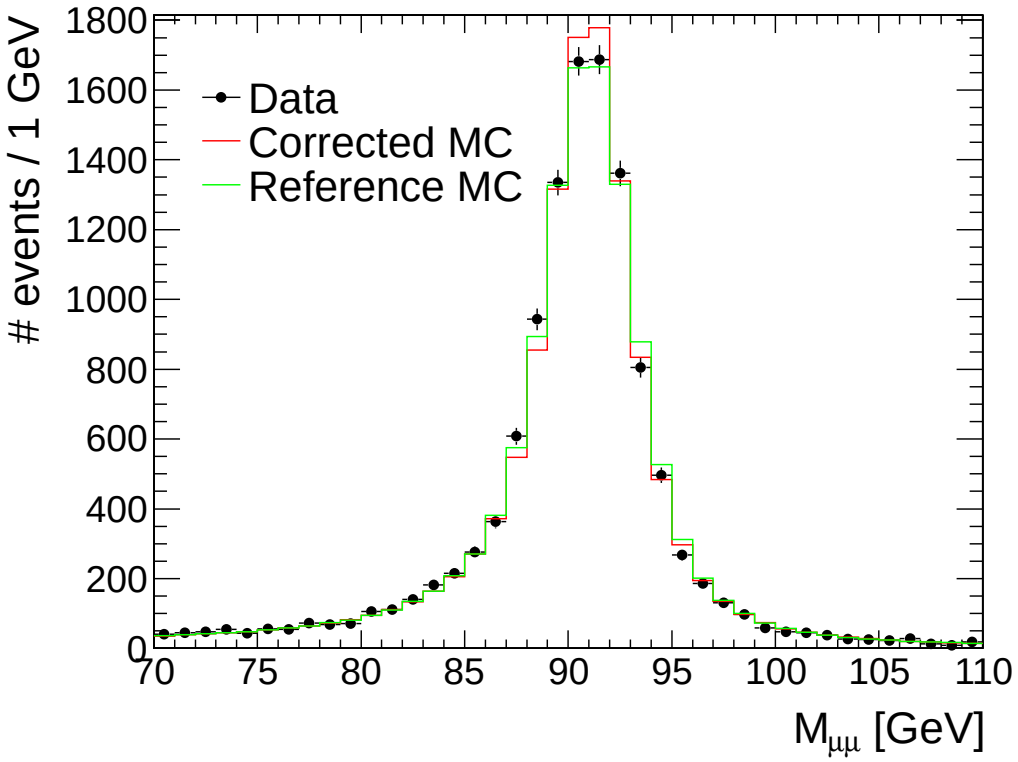

Figure 4.26: Corrected MC - MC distribution after the shift correction. The invariant mass distribution in MC before the correction (green) and MC after the correction (red) are plotted. Experimental points (black dots) are also included for comparison. Note that before the correction the MC was wider than the data. Now it is no longer the case.

muons; the amplitude of the sinusoidal is similar for both charges; the phase of the sinus for positive and negative muons differs in π. All these features already appeared when studying the reduced weak mode. For this reason a parametrization of the form

$$\delta(1/p_T) = constant + A \cdot q \cdot \sin(\phi + \phi_0) \tag{4.8}$$

is appropriate.

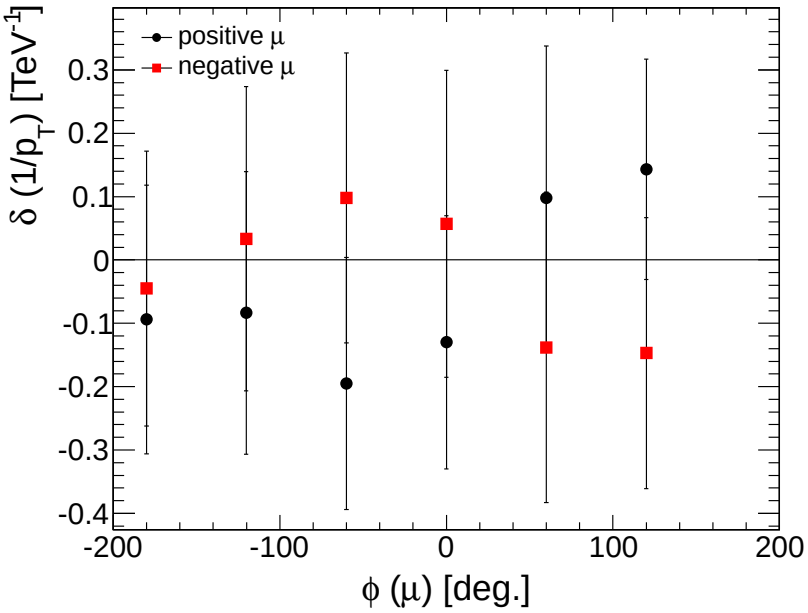

Figure 4.27: Scale factor dependence on ϕ - Scale factor as a function of ϕ for positive and negative muons where the 12 parameters of the fit are left free.

Concerning the dependence of the scale factor on the muon pseudorapidity we first look for a possible dependence on charge. We let the scale factor in η take different values for positive and negative muons. In Fig. 4.28 the fitted values of the scale factor in the different regions in η are plotted. No dependence on charge is observed.

In addition, a possible asymmetry between the positive and negative parts of CMS as a function of the z coordinate is considered. In Fig. 4.29 the variation of the scale factor with η for the whole pseudorapidity range is presented. In this case the most appropriate functional form to parametrize the scale factor is not obvious. Long tails observed in high η regions (the performance of the detector degrades at high η values, due partially to higher track density, giving rise to high residuals) prevent the result

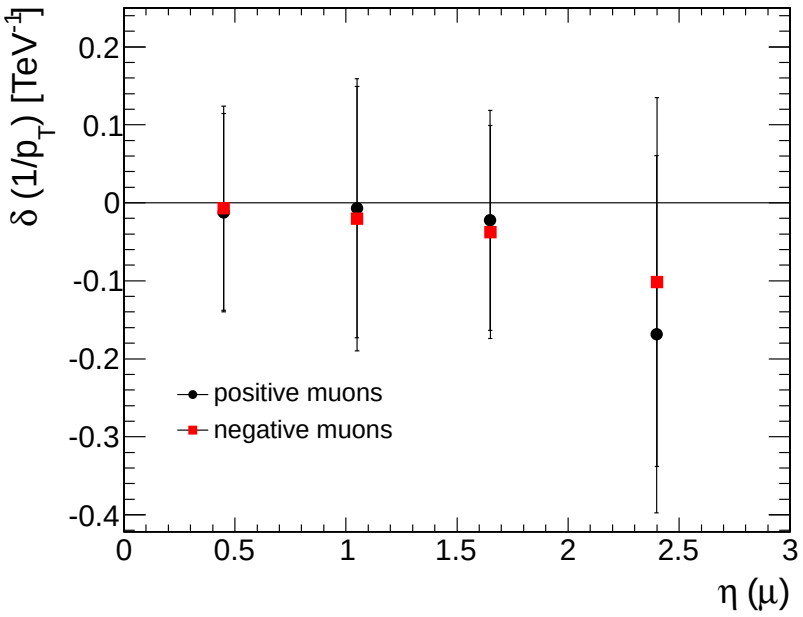

Figure 4.28: Scale factor dependence on η - Study of the dependence of the scale factor term on η and charge. The values of the scale factor for negative and positive muons are similar. The scale factor parameters in each region are allowed to vary freely.

from being highly accurate in that region. Additionally, the asymmetry in the location of the CMS shaft (a vertical tunnel built to descend the different detector components during CMS installation in the cavern) makes one half of CMS to be better aligned that the other one (the region where the shaft receives a bigger amount of cosmic muons which are used for the alignment of the detector). However, a parabolic shape independent with charge seems to be a good compromise between simplicity in the parametrization and a good modeling of the detector behavior. A parametrization of the form

$$\delta(1/p_T) = constant + A \cdot \eta^2 \tag{4.9}$$

is chosen for the variation of the scale factor in η.

4.5.2.2 Resolution dependence in η

In this subsection we analyze the possible dependence of the resolution term on η. The same procedure applied to the scale factor is used in this case. We distinguish 4 regions

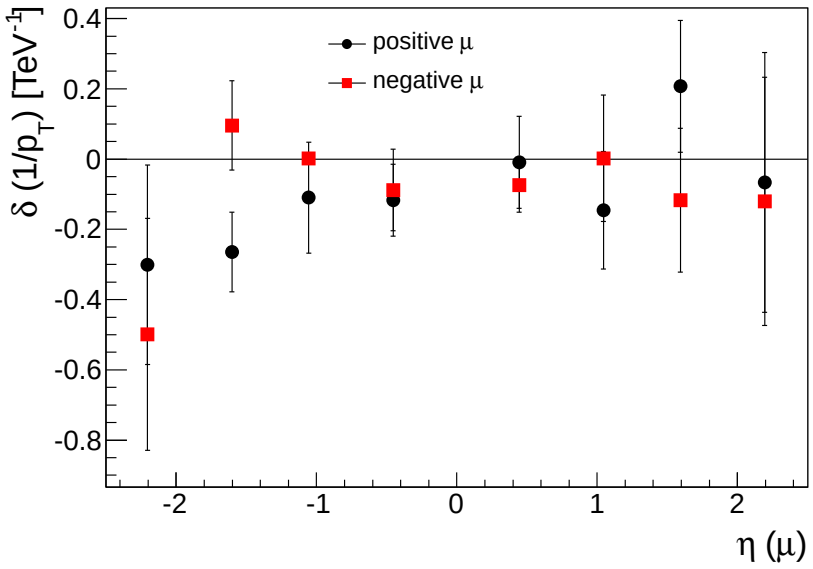

Figure 4.29: Scale factor dependence on η - Scale factor form as a function η for positive and negative muons where the 16 parameters of the fit are free.

in η as before: barrel, overlap region, near endcap and far endcap. By symmetry reasons, we will assume that the resolution term does not depend on ϕ.

In Fig. 4.30 the variation of the resolution with η for the whole CMS is presented. The behavior is symmetric in η. A parabolic shape seems again to be a good choice. We use the form:

$$\sigma(1/p_T) = constant + A \cdot \eta^2 \tag{4.10}$$

for the resolution term.

4.5.3 Scale factor and resolution results after the parametrization

As it has been presented before, the parameters $\delta(1/p_T)$ and $\sigma(1/p_T)$ that control momentum scale and resolution effects, depend on the muon charge, its azimuthal angle ϕ and its pseudo-rapidity η. Once we have defined what is the general dependence of these effects on charge, ϕ and η we will determine the precise value of these effects.

The basis of the method is to modify the reconstructed transverse momentum, p_T, according to the following expression:

$$\frac{1}{p_T} = \frac{1}{p_{T,sim}} + \delta(q,\phi,\eta) + \sigma(q,\phi,\eta)\ \text{Gauss}(0,1) \tag{4.11}$$

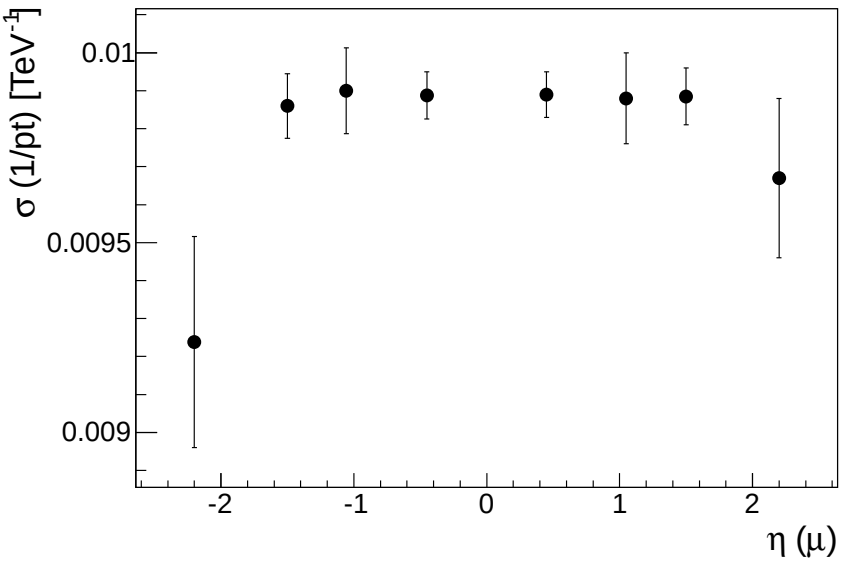

Figure 4.30: Resolution in η - Dependence of the resolution term on η for the whole pseudorapidity range. A symmetric behavior is observed in η. A parabolic function seems to be an appropriate description of the resolution in η.

where $p_{T,sim}$ is the simulated value of the muon p_T, Gauss$(0,1)$ denotes a sampling following a Gaussian function of zero mean and unit variance and δ and σ the parameters controlling the additional effects we want to model. This ansatz is adequate for cases presenting slightly worse resolutions in data than in simulation, and therefore well suited for the present study. According to the results of the previous sections, we can assume angular dependences of the type:

$$\delta(q,\phi,\eta) = A + B\eta^2 + q\ C\ \sin(\phi - \phi_0) \tag{4.12}$$

$$\sigma(q,\phi,\eta) = A' + B'\eta^2 \tag{4.13}$$

where A, ϕ_0, B, A' and B' are the final parameters to be determined.

The result is shown in Fig. 4.31. An excellent agreement for the reconstructed Z invariant mass distribution between data and corrected simulation is reached.

The resulting shifts δ as a function of ϕ and η are shown in Fig. 4.32. Although both data and simulation have an azimuthal bias the phases and amplitudes of the two are significantly different. There is also a significant difference in the η biases where

data shows a bigger effect. Fig. 4.33 shows the transverse momentum resolution versus η after the biases are corrected.

The resolution is defined as:

$$\sigma = \frac{1/p_{T,rec} - 1/p_{T,gen}}{1/p_{T,gen}}$$

The resolution in the reference MC (STARTUP MC) is shown in black dots and the values from the MC, once it is adjusted to reproduce the data, is shown as red squares. A general good agreement is observed. We can also appreciate that the reference MC reproduces well the behavior observed in data for the barrel region ($\eta < 1$). However, in the endcap region, the reference MC overestimates the resolution, and in fact, data exhibit better muon p_T resolution.

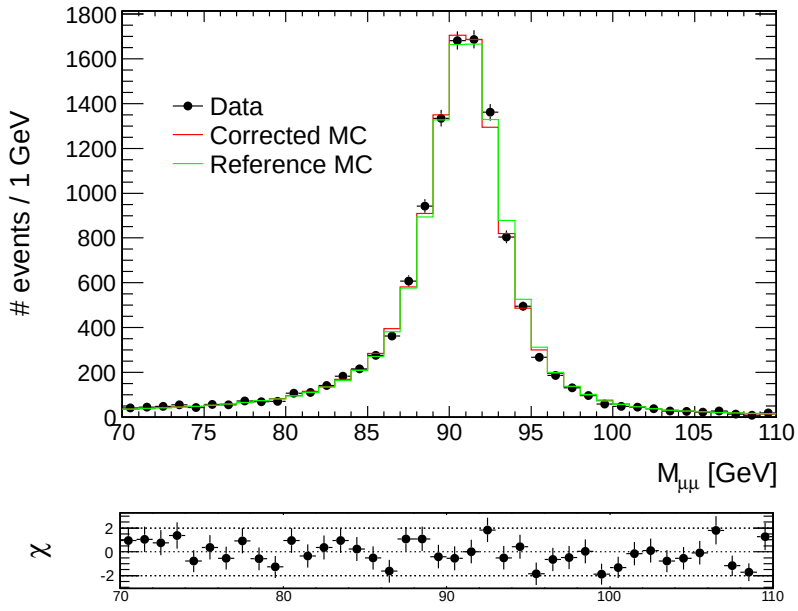

Figure 4.31: Reconstructed Z invariant mass distributions in data and Monte Carlo. The original Monte Carlo prediction (green line) has been corrected (red line) by the parameters determined in the fit.

The fitted values of the parameters are shown in Tab. 4.6.

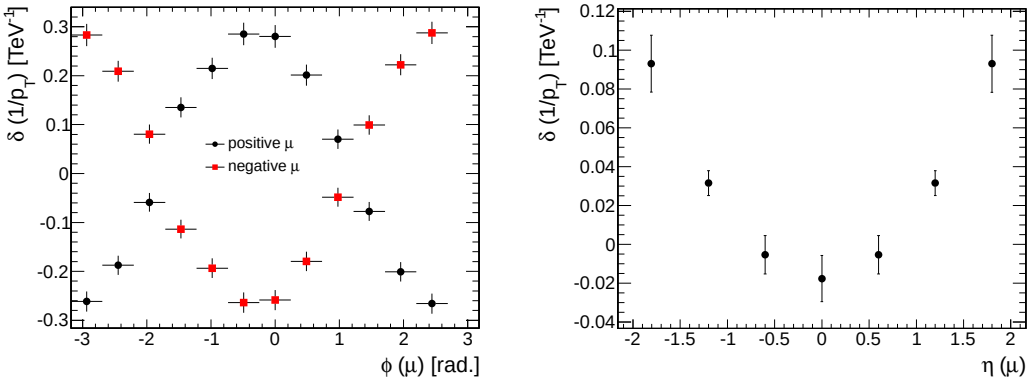

Figure 4.32: Scale shifts obtained as a function of the azimuthal angle for positive and negative charged muons (left) and of the pseudo-rapidity (right).

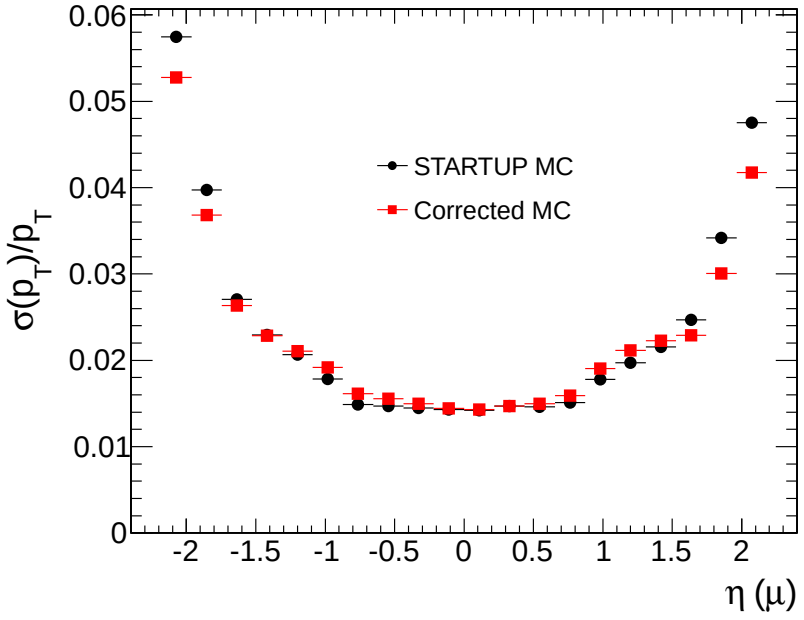

Figure 4.33: Transverse momentum resolution as a function of the pseudo-rapidity of the muon. The black dots represent the resolution from the reference MC and in red the resolution once the MC has been corrected with additional scale and resolution terms.

Table 4.6: Fitted parameters that control muon momentum scale and resolution effects.

Effect	Values
$\delta(q,\phi,\eta) = A + B\eta^2 + q\ C\ \sin(\phi - \phi_0)$	$A = 0.016\ \pm\ 0.019\ \text{TeV}^{-1}$ $B = 0.005\ \pm\ 0.007\ \text{TeV}^{-1}$ $C = 0.282\ \pm\ 0.01\ \text{TeV}^{-1}$ $\phi_0 = -0.134\ \pm\ 0.03$
$\sigma(q,\phi,\eta) = A' + B'\eta^2$	$A' = 0.236\ \pm\ 0.01\ \text{TeV}^{-1}$ $B' = -0.135\ \pm\ 0.005\ \text{TeV}^{-1}$

4.5.4 Comparison with MuScleFit method

The results obtained in the previous section are now compared with those obtained with an independent method (the MuScleFit method), also available in CMS (48).

The MuScleFit method is an alternative method to compute the resolution and scale factors in data. It is based on a multidimensional likelihood fit of a functional form of the Z boson peak given by MC at generator level.

The approach in this method can be summarized in the following steps: ansatz functions are used to describe the muon p_T for scale and resolution, of the form: $p_T = F(x_i; a_j) \cdot p_T$ and $\sigma_i(x_i) = G_i(x_m; a_j)$, with $x_i = \eta, \phi, p_T$ and a_j being the parameters to be computed. A Lorentzian convoluted with a gaussian is the ansatz function used to model the Z boson mass peak. The background is described with a sum of linear and exponential functions. This ansatz function is used to describe the Z boson mass peak. The likelihood of these ansatz functions to reproduce the reconstructed dimuon mas distribution is computed. Note that for this method the resolution and scale factor terms are computed as a whole, and not as extra terms to be added to the already simulated one (as SIDRA does).

A comparison of the result from both methods when applied to data for the scale factor term in ϕ, in terms of the $\delta(1p_T)$ extra scale factor, is shown in Fig. 4.34. A good agreement is observed between both methods. In η the agreement between both methods is also remarkable (see Fig. 4.35).

In Fig. 4.36 a comparison of the results for the resolution term with SIDRA and MuScleFit. Differences from both methods are within statistical errors.

This approach is shown to perform accurately to get the resolution and bias terms. However the method proposed in this thesis presents several advantages: using the simulation as starting point avoids ad-hoc parametrizations of the reconstructed mass at generation level, which are necessarily cumbersome due to the presence of analysis cuts and other theoretical effects (QED final state radiation, PDFs, ...); Monte Carlo generators like PYTHIA (49), POWHEG, etc. are already precise enough in the predictions for the Z invariant mass shape, particularly around the Z pole, and this shape is minimally modified by higher order QCD and QED effects; the simulation already includes inhomogeneities and tails in the response of the detector, which are hard to reproduce with a fully analytical ansatz. In addition, it gives directly the distortion to be incorporated in the MC to fit the data. From a practical point of view this last feature makes this method specially appealing for physics analysis.

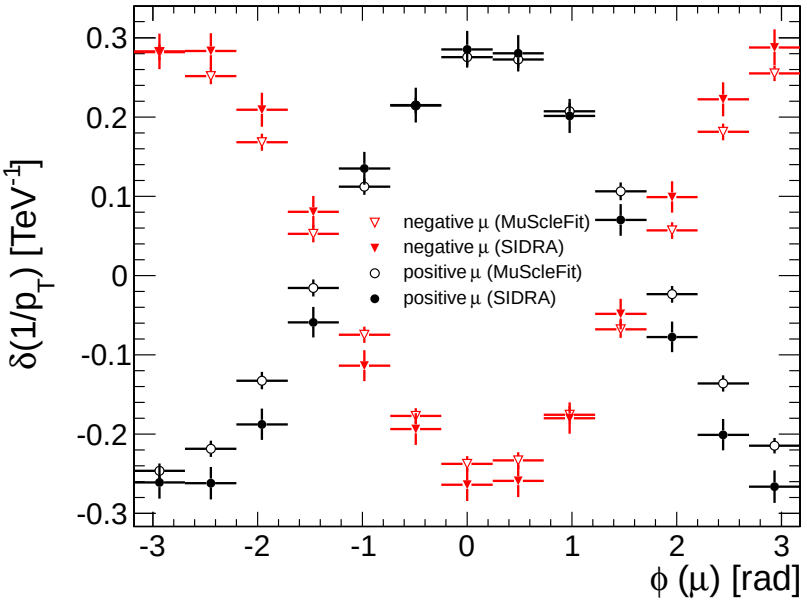

Figure 4.34: Sidra-MuScleFit comparison for the scale factor dependence on ϕ.

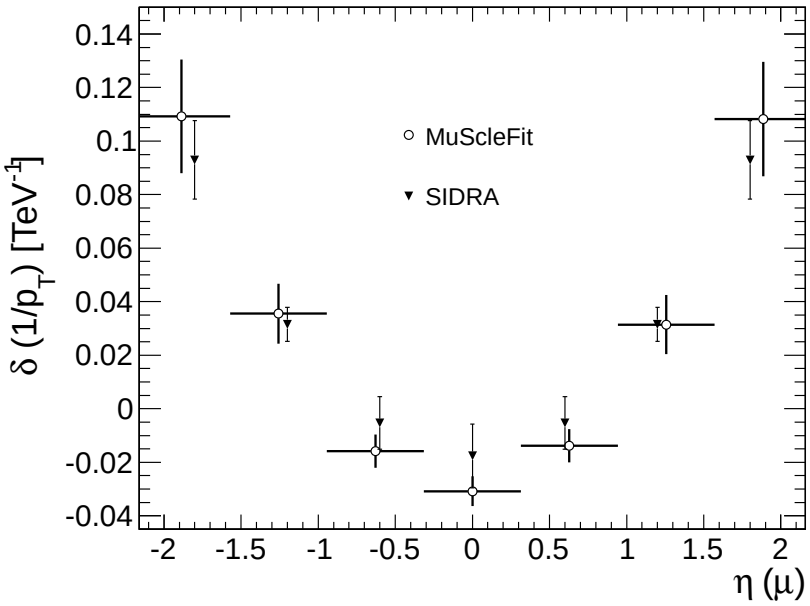

Figure 4.35: Sidra-MuScleFit comparison for the scale factor dependence on η.

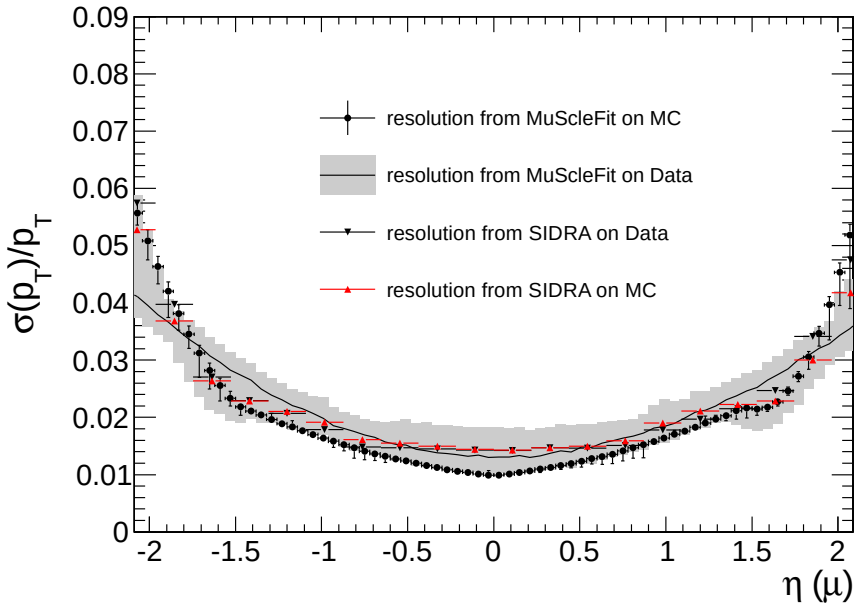

Figure 4.36: Sidra-MuScleFit comparison for the resolution term dependence on η.

4.6 Theoretical uncertainties and biases in the determination of muon resolution parameters

The method described in this Chapter is based on the assumption that the modeling of the $Z \rightarrow \mu^+\mu^-$ invariant mass distribution at the generator level is largely independent on theoretical assumptions. In this Section we discuss how much theoretical uncertainties affect the description of the Z mass peak and therefore propagate to the muon momentum scale and resolution determination.

We will consider PDF uncertainties, ISR and FSR uncertainties, and weak and QED interference effects.

We will obtain an estimate of: a) how much the average value of the invariant mass shifts and b) how much the width of the invariant mass changes. The results from a) can be interpreted as a systematic uncertainty on a scale shift at the Z peak, whereas results from b) can be interpreted as systematic uncertainties on the smearing at the Z peak. In order to ensure that results are meaningful, we will determine these shifts and widths on a simulated and reconstructed $Z \rightarrow \mu^+\mu^-$ sample with a reconstructed dimuon invariant mass cut of $M_{\mu\mu} > 60$ GeV.

Our Monte Carlo sample consists of 55003 $Z \rightarrow \mu^+\mu^-$ selected events. They have been generated with POWHEG (42), which should provide a much more reliable description of the process than PYTHIA, particularly for the transverse momentum distribution of the Z boson, as discussed below. All systematic effects are collected in Table 4.7. From the technical point of view the values of the associated uncertainties will be evaluated with a reweighting procedure. Every element is given a weight depending on the ratio of the reconstructed distribution once one of the parameters of the generation is modified according to its uncertainty, and the original distribution.

4.6.1 Parton density function (PDF) uncertainties

PDF uncertainties can obviously modify the relative fraction of on-peak and off-peak events, and therefore affect the determined parameters. The logic employed to assign uncertainties is the one reported in (50). For the CTEQ6M NLO PDFs (51) used in the POWHEG generation, we observe a possible scale shift uncertainty of $\Delta = ^{+37\pm9}_{-32\pm5}$ MeV. Using MSTW2008 NLO PDFs (45) instead, we do not observe any

change in the central values with respect to CTEQ6M, while the shift uncertainty is reduced to $\Delta = ^{+26\pm16}_{-20\pm15}$ MeV, i.e. a negligible relative shift of 0.03% at the Z peak.

The change in the observed width of the resonance is however larger: $\sigma = ^{+268\pm3}_{-215\pm3}$ MeV with CTEQ6M and $\sigma = ^{+193\pm11}_{-131\pm11}$ MeV with MSTW2008, i.e. a relative smearing of $0.2 - 0.3\%$ at the Z peak, to be compared with the typical ideal resolutions of order $1 - 2\%$. The CTEQ6M numbers are reported in Table 4.7. One has to notice that the MC sample used in the 2010 analysis was simulated with the CT10 PDF set, so the CTEQ6M numbers are very conservative upper values.

4.6.2 Initial State Radiation (ISR) uncertainties and higher order QCD effects

ISR uncertainties are mostly affecting the p_T distribution of the Z boson, and induce an effect very similar to a smearing of the muon p_T. It is evaluated by reweighting the PYTHIA Z boson p_T spectrum obtained with PARP(64)=0.2 (default) with the PYTHIA Z boson p_T spectrum using PARP(64)=0.1, which is known to give a better agreement to the predictions from programs like RESBOS (52) or POWHEG itself. This leads to a scale shift of $\Delta = 62 \pm 41$ MeV, and an extra smearing of $\sigma = 293 \pm 29$ MeV.

Note however that, since we are using POWHEG, this extra smearing is largely overestimated, and we should expect values that are much smaller than the original estimate, i.e. something probably below 100 MeV for σ. Nevertheless, and in order to cover possible higher order QCD effects, ISR numbers reported in Table 4.7 correspond to the most pessimistic case in which we fully rely on a classic PYTHIA generation.

4.6.3 Weak and QED interference effects

Weak corrections are estimated by using a similar reweighting strategy. The essential effect at the Z peak with respect to the Born-improved approach employed in PYTHIA is an increase of the pure Z exchange diagram by 0.5% relative to the photonic one. The result is a negligible additional scale shift and a decrease of the Z width of $\sigma = -18 \pm 1$ MeV. QED interference effects are totally negligible at the Z resonance, and have just some minimal effect at the edge of the invariant mass region (< 70 GeV), where the sensitivity of the method is almost negligible, since shifts and degradations of resolution predominantly manifest very close to the maximum of the Z peak.

4.6.4 Final State Radiation (FSR) uncertainties

FSR QED effects significantly distort the invariant mass shape with respect to the original generator distribution in the absence of QED radiation. However, the dominant effects are reliably reproduced by the parton-shower approach used and PYTHIA (and therefore in POWHEG), since it must correctly reproduce leading-order collinear effects. Possible missing effects are of order $\mathcal{O}(\alpha/\pi)$, i.e. $0.2-0.3\%$ and mostly affecting non-collinear directions with respect to the final lepton. We have tried to estimate missing effects of this order and higher orders by changing the scale of the QED radiation from $Q^2 = 0$ to $Q^2 = p_T^2$, where p_T is the relative photon p_T with respect to the final lepton. This change induces negligible effects: a scale shift of $\Delta = -12 \pm 2$ MeV and an extra smearing of $\sigma = 8 \pm 2$ MeV.

4.6.5 Total systematics

All different sources of systematics and their (rounded) sum in quadrature are compiled in Table 4.7. Overall, theoretical uncertainties can at most produce scale shifts below 0.1% and extra smearings below 0.5%, relative to the value of the Z mass (91.2 GeV). A relative shift of 0.1% in the mass induces a fake momentum shift of order $0.1\% \cdot p_T/[91\text{ GeV}]$ if the effect is assumed to equally affect positive and negative muons. An extra smearing of 0.5% can be interpreted as an extra smearing of $0.5\%/\sqrt{2}$ for muons with $p_T \sim M_Z/2$ GeV, which is smaller or of the order of the ideal tracker resolution in this range.

Note in any case that all these sources of uncertainty will equally affect any other method trying to estimate resolution parameters, since in general they will be hardly distinguishable from authentic resolution effects.

Source	Δ[MeV]	σ[MeV]
PDFs	37 ± 9	268 ± 3
ISR and NLO effects	62 ± 41	293 ± 29
WEAK effects	0	-18 ± 1
FSR effects	-12 ± 2	8 ± 2
TOTAL	73 ± 37	398 ± 22

Table 4.7: Summary of systematic uncertainties considered for resolution parameters. Δ denotes a systematic scale shift, whereas σ denotes and additional increase (or decrease if sign is negative) in the observed width of the dimuon invariant mass. Quoted PDF uncertainties correspond to the size of the maximum observed deviations in the CTEQ6M case, irrespective of the sign (CTEQ uncertainties are asymmetric). The total uncertainty corresponds to a rounded sum in quadrature of all effects.

5

Measurement of the inclusive W boson production cross section in the muon channel

W and Z bosons are copiously produced at LHC proton-proton collisions, being one of the dominant process (53). Although being particles widely studied in past experiments, their importance in the initial phases of LHC operations is clear. When decaying to leptons, they are relevant processes for several reasons: it is a benchmark for lepton reconstruction and identification necessary for other analysis, a precision test of perturbative QCD and the parton distribution functions (PDFs) of the proton, a possible estimator of the integrated luminosity at LHC collisions and a dominant background for many interesting physical searches.

Lepton reconstruction and identification is capital for many analyses. Since all analyses based on muons are clean they are the golden channel of many searches. A correct muon reconstruction and identification is, as a consequence, of paramount importance. They are also used to test and calibrate the detector response in terms of trigger, efficiencies and resolution. The measurement of muon momentum resolution presented in the previous Chapter computed using the Z boson resonance is an example of how the study of electroweak processes can influence the understanding of results in other analyses.

Several electroweak channels give direct access to the proton PDFs. Predictions from different PDFs are sometimes in slight disagreement (54), or their knowledge

is insufficient. A better understanding of certain electroweak processes could help to solve these contradictions and shed light on the quark composition of protons. As an example, the $W + c$ channel that is described in Chapter 6 gives direct access to the strange quark content. The strange quark composition in protons is not well understood and the different PDF groups would benefit for such a study. It will also have influence on important measurements as the W boson mass (55), decreasing the systematic error associated to such a measurement.

The total luminosity delivered by the accelerator is the dominant uncertainty of many analyses (of the order of 4%). The number of W or Z reconstructed bosons is a reliable and stable estimator of the delivered luminosity that could be used in the future.

Electroweak processes can help to understand the signal of important physical searches (for example the Higgs decay in two W bosons) and are background of many others (being even the dominant background in certain cases). A better understanding of electroweak processes allows us to improve our simulations and as a consequence to tune the analyses based on these channels. Take for instance the case of the W'$\rightarrow e\nu$ search shown if Fig. 5.1. As it can be seen in this figure, the dominant background corresponds to electroweak processes (W+jets).

CMS deploys a variety of different studies centered on the production of electroweak bosons in proton-proton collisions. Several electroweak studies are already published (or accepted for publication). These studies are classified:

1. Production of W and Z bosons in pp collisions.
2. Diboson production cross section and constraints on anomalous triple (and quartic) gauge couplings
3. Measurement of electroweak parameters: $\sin^2\theta_W$ and the W mass

The measurement of the **inclusive W and Z boson cross section and cross sections ratios** was the first electroweak measurement performed in CMS (57). It serves as a test of the Standard Model at this new energy. The measurement of the W boson production cross section in the muonic channel is the subject of this Chapter.

W charge asymmetries are also studied (58). W^+ are created via a $u\bar{d}$ current interaction, while in the case of W^- we need a d quark in the initial state ($d\bar{u}$ current

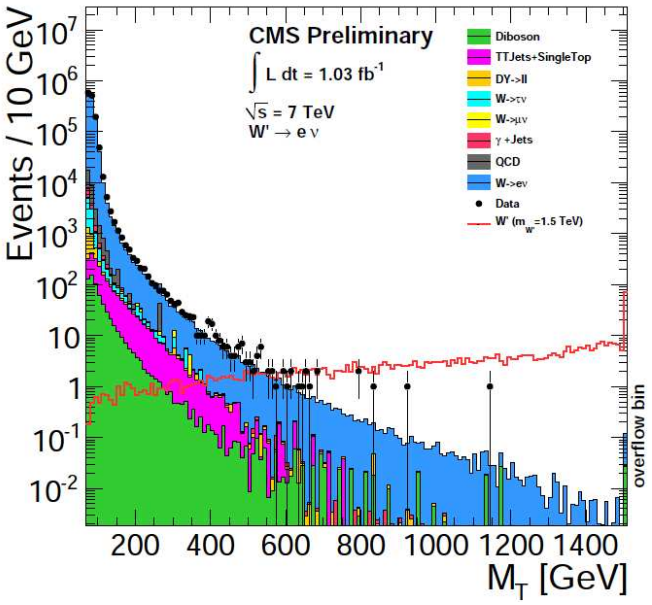

Figure 5.1: - Transverse mass distribution for the W' search in CMS. The $W \rightarrow e\nu$ process is the dominant background in this search (56).

interaction). The relative content of u quarks in the proton respect to d quarks is 2:1, so a W charge asymmetry of around 2 is expected in LHC collisions at 7 TeV. Precision measurement on the W charge asymmetry as a function of the boson rapidity can provide new insights on the u/d ratio and sea antiquark densities in the ranges of the Bjorken parameter x (59) explored at the LHC.

The **differential cross-sections of Drell-Yan pairs** measurement (60) in mass also helps to better understand the parton PDF. This process is theoretically well understood and differential cross sections predictions at the level of NNLO with respect to the inclusive production are available. Differential cross section measurements are thus sensitive to variations due to the imperfect knowledge of the parton PDFs, due to uncertainties in α_S, and the choice of QCD renormalization and factorization scales.

The study of the **polarization of W and Z bosons produced with high transverse momentum (61)** provides a stringent test of perturbative QCD calculations. It is also an important prerequisite to searches for new physics. In this case we look for W and Z boson produced with hard jets. In this situation the W boson will exhibit a high transverse momentum. Due to the V - A nature of the coupling of the W boson to fermions and the enhancement of the gluon-quark contribution to the W+jet pro-

duction with respect to Tevatron collisions we expect W boson with high p_T to exhibit a sizable transverse polarization.

The measurement of the **production of jets in association with W and Z bosons (62)** provides a stringent and important test of perturbative QCD calculations. Predictions at the Next-to-leading order (NLO) level respect to the inclusive production are available for V + n jets, with n up to four for the W and three for the Z (63, 64, 65, 66). The precision of the prediction varies from 10% up to 30% due to uncertainties on parton distribution functions and on the perturbative nature of the calculations. This study is specially relevant since, as shown in the introduction, they constitute an important background in searches for new physics and for studies of the top quark. As a consequence, precise measurement of the cross section and an understanding of the jet and lepton kinematics is essential.

The **production of heavy quark jets in association with W and Z bosons** is the subject of other interesting studies at CMS (67, 68). The production of the Z boson in association with b-quarks is an important measurement at the LHC. It represents a benchmark channel to the production of the Higgs boson in association with b-quarks. At the same time it is an important Standard Model background to many new physics searches and Higgs in final states with leptons and b-jets. The production of W bosons in association with a c-quark gives direct access to the s-quark content of the protons. A better understanding of this process can improve the knowledge of the proton PDFs on the s-quark. This study is the subject of next Chapter.

Electroweak boson production with tau leptons in the final state are an important probe for many new physics processes in proton-proton collisions at the LHC. Both electroweak boson decays into taus ($Z \rightarrow \tau^+\tau^-$ and $W \rightarrow \tau\nu$) are studied within the CMS collaboration (69, 70). Tau decays can be purely leptonic (in electrons, muons and neutrinos) or semileptonic (neutrino plus hadronic system). Both possible decays are studied by the CMS electroweak group. Tau channels can be signal of many interesting studies. The study of the decay in the semileptonic channel is essential for searches for new physics based on tau leptons. Tau leptons can also be an important feature in signals for supersymmetry and extra dimensions, and in searches for extra gauge bosons. Taus are also an important background of relevant processes. As an example, the $W \rightarrow \tau\nu$ production is the major background in the search for the charged

Higgs boson in the $\tau\nu$ final state. The $W \rightarrow \tau\nu$ production has to be well understood as a test of the Standard Model.

Another electroweak study in CMS is the **forward-backward asymmetry of Drell-Yan pairs (71)**. This study is interesting for a host of reasons. The process $q\bar{q} \rightarrow Z\gamma^* \rightarrow l^+l^-$ involves both vector and axial-vector couplings of electroweak bosons to fermions, resulting in a forward-backward asymmetry in the number of lepton pairs. This asymmetry depends on the dilepton invariant mass, quark flavour and the electroweak mixing angle. So deviations from the Standard Model prediction for the asymmetry may indicate the existence of a new neutral gauge boson, quark-lepton compositeness, supersymmetric particles, or extra dimensions. This measurement can also contribute to improve QCD measurements with high order corrections and to constraint parton PDFs. The electroweak mixing angle can also be measured using this asymmetry. Finally, the asymmetry as a function of the mass also provides information on the u and d quarks separately.

Diboson processes are also studied at CMS. The self-interaction between gauge bosons occurs via the trilinear gauge boson couplings (TGCs) and it is well understood within the Standard Model. Boson self-interactions are a consequence of the non-abelian $SU(2) \times SU(1)$ group which describes the electroweak force within the Standard Model. Their couplings are predicted by the gauge structure of the Lagrangian. The high sensitivity of the production of $Z\gamma$, $W\gamma$, WW, WZ and ZZ bosons in proton-proton collisions to these TGCs makes these channels to be of special interest to test the Standard Model (72, 73). Any deviation observed from the SM prediction would be a hint of new physics (new particles decaying to a diboson state or a new interaction that would modify the strength of the couplings).

The $\sin^2\theta_W$ is the only free parameter in the Standard Model that fixes the relative couplings of all fermions to the neutral weak bosons (74). The weak-mixing angle θ_W describes the rotation that transform the original W_0 and B_0 vector bosons states into the observed γ and Z bosons. This measurement is of special interest to test the universality of the fermion-gauge boson couplings and predictions of the Standard Model. Precision measurements at LHC aim to reach the level of 1% accuracy in this parameter.

The measurement of the **mass of the W boson** presents several technical difficulties that can't be addressed until the integrated luminosity recorded in CMS is of the

order of 10 fb^{-1}. The mass of the W boson is known at the level of 20 MeV by the combination of measurements in different experiments. With the current precision, the combined measurement top mass-W boson mass is one sigma away from the Standard Model prediction.

5.1 W and Z boson production in CMS

At high energy, a proton-proton collision can be interpreted at the hard scattering level as a collision of either quark-quark, quark-antiquark, gluon-quark or gluon-antiquark partons. The higher the energy of the primary protons, the bigger the radiation and hence the influence of gluons and sea quarks. At the LHC energy, W and Z bosons will be produced by interaction of all flavour of quarks and antiquarks. Z bosons are dominantly produced via u $\bar{u}$ or d $\bar{d}$ annihilation. For the W production, the dominant processes are $u\bar{d} \rightarrow W^+$ and $d\bar{u} \rightarrow W^-$. The LHC is a W and Z factory, with over 100 bosons produced per second in nominal conditions.

The total cross-section of a certain process $pp \rightarrow X$ at a center of mass energy $\sqrt{s}$ can be written (75):

$$\sigma = \sum_{i,j} \int \hat{\sigma}_{ij}(\hat{s}, \mu_f, \mu_r) \int_0^1 \int_0^1 f_i(x_1, \mu_f^2) f_j(x_2, \mu_f^2) \delta(\hat{s} - x_1 x_2 s) dx_1 dx_2 d\hat{s} \tag{5.1}$$

where $\hat{\sigma}_{ij}$ are the parton-parton cross-section $i + j \rightarrow X$, $\sqrt{\hat{s}}$ the reduced center-of-mass energy of the parton reaction, x_1 the energy fraction of parton i, x_2 the energy fraction of parton j, $f_i(x)$ the parton distribution function for parton i (same for j), μ_f the factorization scale and μ_r the renormalization scale. The partons i and j may be the valence quarks of the proton (u,d), the gluons interchanged by the quarks (g) or any of the other sea quarks present in the proton ($\bar{u}$,$\bar{d}$,s,c,b,t,$\bar{s}$,$\bar{c}$,$\bar{b}$,$\bar{t}$). The delta function considers the energy conservation of the process.

In a very simplified case of the production of a narrow resonance, the previous cross-section can be written as

$$\hat{\sigma}_{ij}(\hat{s}) = \sigma_{ij} \delta(\hat{s} - M^2) M^2 \tag{5.2}$$

with M being the mass of the resonance and σ_{ij} the constant cross-section of the reaction $i + j \rightarrow X$ at the peak of the resonance. Once applying this simplification, the total

cross-section of the process $pp \rightarrow X$ can be written as

$$\sigma = \sum_{i,j} \sigma_{ij} M^2 \times L_{ij} \tag{5.3}$$

being L_{ij} the parton-parton luminosity

$$L_{ij} = \frac{1}{s} \int_{\frac{M^2}{s}}^{1} \frac{1}{x} f_i(x) f_j(M^2/xs) dx \tag{5.4}$$

This simplified picture can be improved by taking W and Z width effects into account, as well as QCD and QED radiative corrections.

Collision data were recorded at center-of-mass energies of $\sqrt{s}$ 7 TeV and 2.76 TeV. W bosons are copiously produced in LHC events due to the high energy available in the collisions. With NNLO cross-sections (including branching fractions to leptons) of 18456 pb for $W^+ \rightarrow l^+\nu_l$ and 12858 pb for $W^- \rightarrow l^-\nu_l$ (where l considers the whole contribution to leptons e, μ or τ) at 7 TeV, the production of W boson decaying into leptons are one of the most important processes that occur at LHC collisions. The production of Z bosons is about 3 times smaller than that of the W bosons.

5.2 W and Z boson decays

W and Z bosons are massive particles that decay rapidly (half-life is of the order of 10^{-25} s). They can only be detected through their final states. Their decays into lepton and quarks are accurately predicted by the gauge theory of electroweak interactions and strong forces.

W bosons can decay either to leptons or quarks. Their leptonic decay can be either $W^\pm \rightarrow e^\pm\nu_e$, $W^\pm \rightarrow \mu^\pm\nu_\mu$ or $W^\pm \rightarrow \tau^\pm\nu_\tau$ via the weak force. W bosons can also decay into quarks following the scheme $W^+ \rightarrow u\bar{d}$ or $W^- \rightarrow d\bar{u}$ where the d quark can be replaced by any d-quark type (s or b quarks). W bosons can not decay into t quarks since the t quark mass is bigger than the W boson mass.

Theoretical predictions for the results shown in the following Chapters are computed at NNLO using FEWZ (Fully Exclusive W and Z production) (63, 76) and the CT10 set of PDFs. With uncertainties of the order of 2% to be reached at LHC, NLO predictions in the strong coupling (order of 10% accurate) are insufficient for a precise comparison with data. A NNLO is necessary in this case.

W boson reconstruction using its decay in quark-antiquark pair, even if it is dominant with respect to the leptonic channels, is humped by the high background coming from QCD interactions. Within the leptonic channels, the muon channel is specially clean. This is the reason why the analysis presented in this thesis are based on the reconstruction of W bosons when decaying into muons.

5.3 W signature in the muon channel

W bosons are produced at LHC almost at rest, with a typical p_T of the order of 4-5 GeV. Therefore, muon and neutrino from its decay will be produced back-to-back in the transverse plane, and the transverse momentum of the muon and neutrino will be relatively high, peaking typically at one half of the W boson mass. The neutrino will not be detected but its presence can be inferred as an apparent lack of energy-momentum balance in the event. This energy-momentum balance can only be evaluated in the transverse plane as most of the particles produced in the collision escape through the beam pipe. By the same reason, the invariant mass of the muon-neutrino system can not be computed since the longitudinal momentum of the neutrino can not be measured, but a partial reconstruction in the transverse plane can be performed, making use of the neutrino and muon variables in this plane.

In Fig. 5.2 a typical W $\rightarrow \mu\nu$ event as recorded by the CMS detector is shown. The characteristics pointed out before are clearly seen in this Fig.: an isolated, high p_T muon and essentially nothing else. The green arrow in the event display indicates the direction of the missing transverse energy (MET) obtained as the negative vector sum of the transverse momentum of all particles in the event. As it can be appreciated in the figure, the direction of the muon and the MET are almost back-to-back.

Several other physical processes occurring on high energy proton-proton collisions also produce events with a similar signature: a big amount of MET and a high p_T muon. The main one are QCD processes. Other electroweak processes like $W \rightarrow \tau\nu$ with the τ promptly decaying into a muon will have a similar signature, although with a softer p_T spectrum. Also Z boson production with the boson decaying into a muon pair should be considered. If one of the muons from the Z boson decay escapes detection either because it lies beyond the active region of the detector ($|\eta| > 2.4$) or due to detector inefficiencies, the event will look like a typical W boson event. Other processes like top

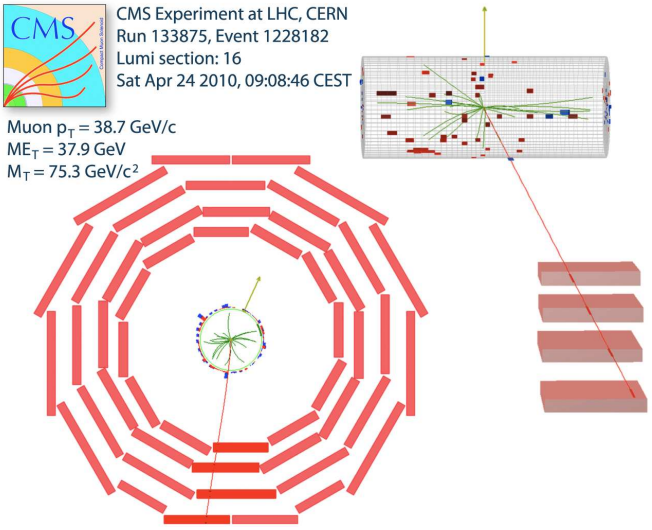

Figure 5.2: $W \to \mu\nu$ **event** - W candidate as detected with CMS detector. The presence of a W boson is inferred by the detection of a high p_T muon and a big amount of MET

production ($t\bar{t}$ or single top) and diboson production (WW, WZ, ZZ) also give rise to events with a similar signature. Compared to the inclusive W production, their cross section is rather small and will not contribute experimentally in a significant amount. However, for the W+jet studies with smaller production cross section, top events (like $t\bar{t}$ and single top events) will be an important background. The top, decaying into a W plus a b quark, has a similar signature to that of the signal.

In Fig. 5.3 the p_T distribution of the muon of a $W^- \to \mu^-\nu$ simulated sample by MC is shown. In the same figure, the p_T distribution of muons coming from QCD simulated processes is also plotted. Muons from W bosons range from 20 to 60 GeV in transverse momentum, with a peak at $\approx M_W/2$. QCD events instead, have a muon p_T spectrum peaking at low values, $p_T \lesssim 20$ GeV. Therefore, setting a threshold at about 20-30 GeV will allow to discriminate between events originated in these two processes. In Fig.5.4 the isolation variable used in the analysis is plotted for the W $\to \mu\nu$ sample and for the QCD events. This variable is defined as the quantity of energy in a cone of 0.3 in ΔR around the muon direction over the p_T of the reconstructed muon,

$$I^{\text{rel}}_{\text{comb}} = \sum (E_T(\text{ECAL}) + E_T(\text{HCAL}) + p_T(\text{tracks}))/p_T(\mu)$$

with ΔR defined as

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

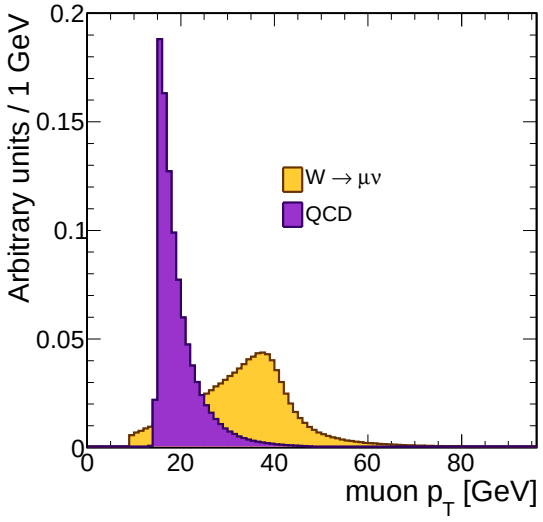

Figure 5.3: p_T distribution - Muon p_T distribution for the $W \to \mu\nu$ channel and QCD processes. The sharp fall at low muon p_T for the QCD sample is due to the generated phase space ($p_T^{\mu} > 15$ GeV). The decrease for the signal (at ≈ 10 GeV) is a consequence of the trigger requirement. The distributions are after the selection detailed in Section 5.4.3.

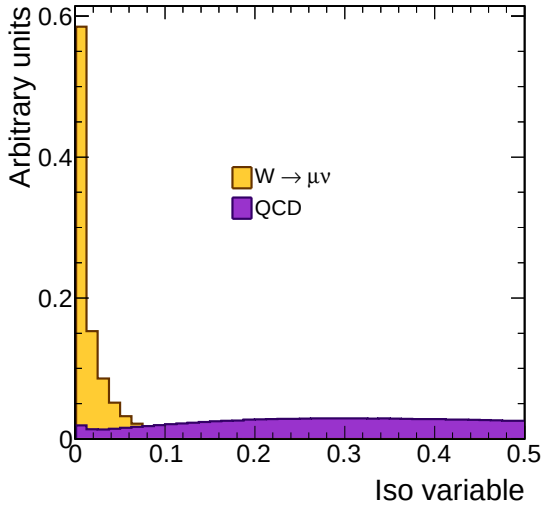

Figure 5.4: Isolation variable distribution - Isolation variable of the muon defined as the quantity of energy (computed using information from the tracker and the calorimeters) in a cone of 0.3 in ΔR around the muon direction over the p_T of the reconstructed muon, for QCD and $W \to \mu\nu$ signal. The distributions are after the selection detailed in Section 5.4.3.

As seen in this Fig.5.4 muons from W boson exhibit a low value of the isolation variable. Muons from QCD events show a more even distribution, spanning a large range of isolation variable. An isolation value of 0.1-0.15 clearly divides a signal region (isolation $<$ 0.1-0.15) and a QCD region (isolation $>$ 0.1-0.15).

The M_T distribution is defined as:

$$M_T = \sqrt{2p_T(\mu)\text{MET}(1 - \cos(\Delta\phi_{\mu\text{MET}}))} \tag{5.5}$$

Fig. 5.5 shows the M_T distribution for $W \rightarrow \mu\nu$ and QCD events as expected from MC. The signal is found at high M_T values (bigger than 40 GeV). QCD events are concentrated at a lower value of M_T.

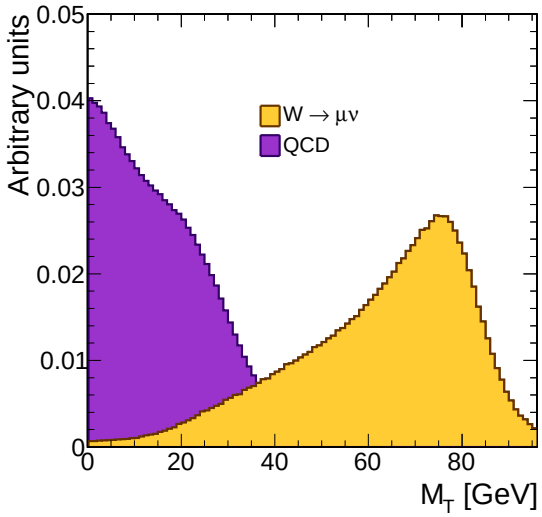

Figure 5.5: M_T distribution - M_T distribution of the muon-MET system for $W \rightarrow \mu\nu$ and QCD events. A requirement of 50 GeV on the transverse mass may be applied to get a pure sample of W bosons. The distributions are after the selection detailed in Section 5.4.3.

In Fig. 5.6 the acoplanarity of the muon-neutrino system defined as:

$$Acoplanarity = \pi - |\phi_1 - \phi_2| \tag{5.6}$$

is plotted for the signal and the QCD background. Signal shows a distribution that corresponds to a system muon-neutrino preferentially back-to-back. QCD is flatter.

In the following sections we will address the precise selection criteria established to extract a clean sample of W boson candidates, based on the variables just presented.

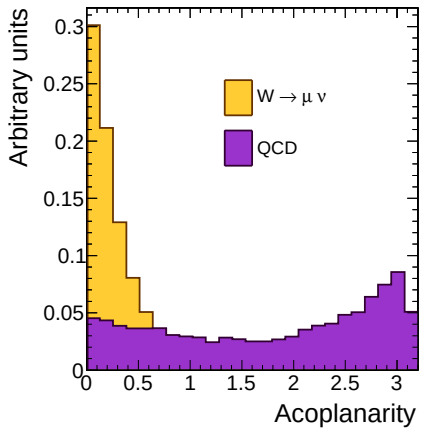

Figure 5.6: Acoplanarity distribution for the signal (yellow) and the QCD contribution (purple) in arbitrary units. Signal events are distributed preferentially at low acoplanarity.

5.4 Inclusive W boson production cross section

CMS published a first measurement (57) of the W inclusive cross section with the first data collected in 2010, corresponding to an integrated luminosity of 3 pb^{-1}. The precision of the measurement was already better than 5% (excluding luminosity uncertainties) being dominated by systematic effects (4%) versus the statistical uncertainty (1%). With the statistics available at the end of 2010 run (10 times bigger) new techniques can be envisaged to reduce the uncertainty associated to some of the ingredients of the measurement and make a high precision determination. In this Section, the strategies followed for a high precision measurement and the results with the data available at the end of 2010 run are shown. The ratio W^+ over W^- is also computed. Many systematic uncertainties are cancelled in the ratio (e.g. the luminosity uncertainty that amounts to some 4% is completly cancelled when computing the ratios), thus the ratio will allow a strict test of the SM prediction.

The number of events produced through a certain physical process observed in the detector after a collision is related to the luminosity provided by the machine and the cross section of the process as:

$$N = \sigma \mathcal{L} A \epsilon \tag{5.7}$$

where σ is the process cross section, $\mathcal{L}$ the integrated luminosity, A the detector ac-

ceptance (fraction of events actually observed in the detectable phase space) and ϵ the detector efficiency (number of candidates after the selection process over the number of total events in the detectable phase space). This expression will also allow us to measure the production cross section of any process by identifying and counting the number of events of that kind registered once we know the luminosity, the acceptance and efficiencies.

What we measure is the total number of events that are reconstructed as W candidates. This sample contains, in general, events coming from other processes with a similar signature as the signal, that is, background events. The acceptance takes into account the fraction of events that escape detection because their final product lies beyond the detector sensitive region. Events that are not reconstructed due to detector inefficiencies or due to the selection requirements are considered with the efficiency term. The cross section is measured as

$$\sigma = \frac{(N_{signal} - N_{background})}{\mathrm{A}\epsilon\mathcal{L}}. \tag{5.8}$$

The signal we are interested in consists of W bosons when they decay on the muonic channel. Thus the signal signature, as explained before, consists of a high p_T, isolated muon plus a high imbalance of the energy measured in the transverse plane. However, events produced by other processes will show a similar signature. They constitute a background for the measurement. As we already introduced in the previous section, we will rely on the signal characteristics to maximize the discrimination between signal and to evaluate the remaining background in the final sample. An effort has been made to minimize reliance on MC modeling, trying to develop methods based on data control sample to control the background, at least the most critical ones. QCD processes with a muon in the final state represent the dominant background as they are copiously produced in proton-proton collisions. Several techniques to reduce and control this background are shown in this Chapter. Events produced by other electroweak processes also present a similar signature: $\mathrm{W} \to \tau\nu$ and $\mathrm{Z} \to \mu\mu$ are the most important ones. Additionally, minor contributions to our background are the diboson production (like WW or WZ events) and $\mathrm{t\bar{t}}$ production. To evaluate their contribution we will rely on the MC simulation for the shape of key distributions as well as the cross section prediction.

In the following sections we will show the data and MC used in the study, the event selection to optimize the W cross section measurement and we will cover the measurement of the different terms contributing to the production cross section.

5.4.1 Data and Monte Carlo samples

In this analysis the full 2010 data set, 36 pb^{-1} of collision data are used. They were collected on-line by the lowest non-prescaled single muon trigger. In order to keep constant the bandwidth of the trigger selection, the trigger threshold was raised every time the instant luminosity increase significantly. At the beginning of data taking, the single muon trigger accepted events with a muon with a p_T bigger than 9 GeV. This threshold was increased and in the end of 2010 run a trigger on muons with p_T bigger than 15 GeV was implemented. In Tab. 5.4.1 the different triggers with their threshold depending on the period of data taking is shown.

Sample	Run range	Int. Lumi pb^{-1}	Trigger Path
Run A	136033-144114	3.07	HLT_Mu9
Run B	146239-148058	4.41	HLT_Mu9
Run B	148059-149442	28.61	HLT_Mu15
Run A+B	136033-149442	36.09	

Table 5.1: Trigger path used for the different periods of data taking.

Once the event is triggered it is centrally reconstructed, and streams of different objects are created to easier the physical analysis. This streams are devoted to a certain physical object (i.e. muons) and are stored in different computing sites, available for the physical analysis.

Even if the information from data is maximally exploited in all the analyses steps, large samples of events simulated with MC are also used in the analysis. They are employed in several steps of the analysis, for cross checks and to asses the validity of the methods deployed.

Simulated samples of signal events and from other electroweak processes (Z boson production and decay, and other W boson decays) are produced using a NLO MC generator (POWHEG) (42, 43) interfaced with PYTHIA (49) for showering and hadronization. $t\bar{t}$ events are generated with PYTHIA. QCD processes with muons in

the final state are also simulated. For both signal and the background several minimum bias events are superimposed to the hard interaction to simulate the effect of the "pile-up" of several proton-proton collisions occurring at the same bunch crossing.

In Tab. 5.2, a detailed table with the MC samples used in the analysis is shown.

Generator	Process	Kinematic cuts (in GeV, $c = 1$)	σ (pb)	Events	PDF set
POWHEG	$W^+ \to \mu^+\nu$	no cuts	5825	~2M	CT10
POWHEG	$W^- \to \mu^-\nu$	no cuts	3954	~2M	CT10
POWHEG	$W^+ \to \tau^+\nu$	no cuts	5825	~2M	CT10
POWHEG	$W^+ \to \tau^-\nu$	no cuts	3954	~2M	CT10
POWHEG	$Z \to \mu\mu$	$m_{\mu\mu} > 20$	1631	~ 2M	CT10
POWHEG	$Z \to \mu\mu$	$20 > m_{\mu\mu} > 10$	3216	~ 2M	CT10
POWHEG	$Z \to \tau\tau$	$m_{\tau^+\tau^-} > 20$	1631	~ 2M	CT10
PYTHIA	$t\bar{t}$	no cuts	162	~1M	CTEQ6L1
PYTHIA	Inclusive μ QCD	$\hat{p}_T > 20$, $p_t(\mu) > 15$, $\lvert\eta_\mu\rvert < 2.5$	84679	25M	CTEQ6L1
PYTHIA	WW	no cuts	42.9	~1M	CTEQ6L1
PYTHIA	WZ	no cuts	18.3	~1M	CTEQ6L1
PYTHIA	ZZ	no cuts	5.9	~1M	CTEQ6L1

Table 5.2: Summary of analyzed Monte Carlo samples for the various signal and background processes. In the first column the generator used to simulate the sample in the second column is shown. Third column shows the generator cuts for the sample production. In forth column the cross section computed at NLO for each process is presented with the number of total events in the fifth column. Last column shows the PDF set used in the generation of the sample.

5.4.2 Muon reconstruction and identification

Typical W boson events consist of a high p_T, isolated muon, with a big amount of MET consistent with the existence of a high E_T neutrino. A good muon reconstruction is required to suppress instrumental background from punch-through, meson decay in flight, cosmic muons, etc. and to ensure an accurate transverse momentum measurement.

Muons are reconstructed separately in the tracker and in the muon chambers. In order to guarantee an excellent reconstruction, the muon has to be reconstructed as a global muon and as a tracker muon (36). For the typical momentum of muons coming from a W boson, the resolution in the transverse momentum of the muon is dominated by the tracker measurement. Thus a minimum number of hits in the tracker when building the muon trajectory is required to guarantee an accurate muon transverse

momentum measurement. At least 10 hits in the tracker and 1 hit in the pixels should take part in the track reconstruction. Also to avoid pathological bad reconstructed muons we reject events with a poor χ^2 of the global muon fit. Events with the muon reconstructed with a χ^2 over the number of degrees of freedom bigger than 10 are not considered.

In order to avoid missidentification of muons (punch-through in our muon chambers) we require the events to be detected by at least two muon chambers. Punch-through particles will lose energy in the iron of CMS wheels and will not reach the second station most of the times. At the same time, this requirement makes the selection to match the requirements used in the trigger. It is also required that at least one hit from the muon chambers participate in the final muon track fit.

Besides the background originating from other physical processes produced in the proton-proton collisions additional contamination may come from cosmic muons crossing the detector. Background from on time cosmic muons (those synchronous with CMS data taking) is reduced by requiring the track to pass within 2 mm to the primary interaction point in terms of the impact parameter in the transverse plane. When applying both requirements (on time muon and small impact parameter) the contamination from cosmic events is negligible. In Fig. 5.7 the impact parameter distribution is shown. All the collision events are present in the central bin. To estimate the remaining number of cosmic muons entering this bin, we extrapolate from the region dominated with cosmics. We assume the rate of cosmic muons arriving to CMS is constant in the impact parameter. The high range of the impact parameter distribution is fitted to a flat distribution. This rate is then extrapolated to the low impact parameter region. It gives an estimation of the number of this kind of events in the sample to be of the order of 1 in 10^4.

5.4.3 $W \to \mu\nu$ selection

We already anticipated the muon criteria to select a pure sample of W boson events: muon p_T bigger than 20 or 25 GeV, isolation variable below 0.1 or 0.15. The other characteristic of this type of events, the presence of a large E_T imbalance, will serve as the final discriminant variable to distinguish between signal and background.

The aim of the selection process is to remove as much background as possible keeping our signal as much as we can. The dominant background are the inclusive

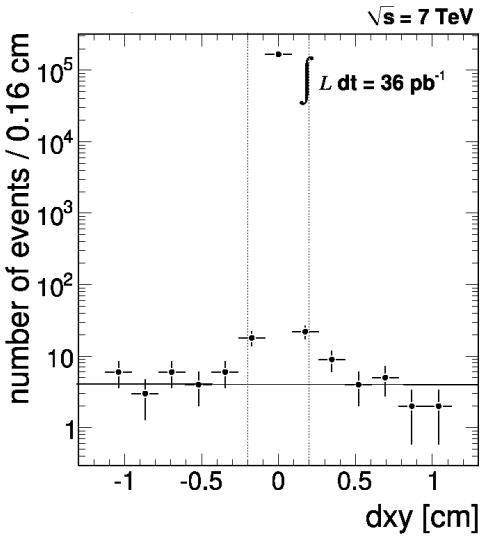

Figure 5.7: Transverse impact parameter distribution. - Distribution of the muon track impact parameter in the transverse plane (referred to the beam spot) of the selected muon, before cutting in this variable. The vertical lines indicate the threshold that define the signal region $|d_{xy}| < 2$ mm.

muons produced as a consequence of QCD interactions. In the first analysis performed in CMS (57) the selection requirements were set rather loose: a muon p_T threshold at 20 GeV and an isolation threshold at 0.15. The compromise was to keep the signal efficiency as high as possible, given the limited size of the sample, at the cost of a relative large QCD background. One of the main contributors to the systematic uncertainty budget was precisely the modeling of the QCD background in terms of the discriminant variable.

The full 2010 data set, ten times larger than the one used in (57), offers the possibility to review the selection process with the aim to reduce mainly the systematic associated to the background subtraction.

The selection optimization cannot focus only on the background subtraction uncertainty, but one has to evaluate the whole picture. Any selection process provoke also a loss of signal efficiency that will have to be corrected later in the analysis. The evaluation of these correction factors induce as well certain systematic uncertainties that will have an impact in the final measurement. In general, the larger the correction factors will be, the larger the systematic uncertainties will be.

To optimize the selection process we will play with the two main variables to discriminate between signal and background: the muon p_T and the isolation variable. The figure of merit we will use in the optimization process is the systematical uncertainty on the cross section measurement.

The sources of systematic uncertainties we will consider are the background modeling, the initial state radiation (ISR) effect and the PDF uncertainty in the evaluation of the signal acceptance. As we remove the QCD background component we expect the background modeling to be less important in the cross section measurement. As a consequence, when performing a more restrictive cut we expect the background related systematical uncertainty to decrease. The ISR effect related systematic uncertainty is expected to increase when making the selection process more restrictive. The typical muon p_T in a W boson decay is between 20 and 50 GeV, with a maximum between 30 and 40 GeV. When moving the muon p_T requirement to higher values, entering the signal region, any change in the modeling of the boson p_T (and the muon one) affects directly the measurement. The uncertainty due to the ISR effect affects the muon p_T, and, as a consequence, the increase of the systematical uncertainty with the increase of the muon p_T requirement is expected.

The complete list of systematical uncertainties considered and the way they are computed will be explained in the following sections. However, a brief description of the way they are computed is necessary to understand this optimization process.

QCD **background subtraction**. Background subtraction is performed evaluating the MET spectrum of the candidate sample. No reliable description of QCD processes is available in the MC and thus an effort was made to derive this shape directly from control samples in data. Several options were considered. First, the MET shape was taken from a QCD-enriched control sample, taking the candidate sample of events with a high-p_T muon, but non-isolated (unlike the signal characteristics). This MET template can be improved, as it will be explained later, to account for observed correlations with the isolation variable. Finally, the MET shape is predicted by the MC, although known not to be fully reliable, is also used. The maximal difference in the W cross section obtained using the different templates is taken as an estimate of the systematic uncertainty associated to the method.

The other two sources of uncertainty considered lie more on the theoretical side as they affect the determination of the correction factor to extrapolate the measurement

in a reduced fiducial volume to the whole phase-space.

In the first case we study how this correction factor varies with the different parameters that govern the modeling in the MC of Initial State Radiation (ISR) effects. This variation is studied for different values of the parameters that define the measurement phase-space: varying the p_T threshold from 20 GeV to 30 GeV, and for two values of the isolation threshold, 0.1 and 0.15.

Similarly we evaluate the impact of the uncertainty on the knowledge of the Parton Distribution Functions (PDFs) in the determination of the acceptance factor, for different values of the p_T threshold, from 20 GeV to 30 GeV, defining the measurement phase-space. For that purpose, several PDF sets are used and variation in the acceptance factor evaluated.

In Fig. 5.8 the three systematic uncertainties studied in this section and their dependence with the p_T requirement is shown for an isolation variable maximum of 0.10 (black dots are used for the QCD background modeling systematic uncertainty, the red squares for the ISR effect one, green triangles for the PDF one and blue triangles to represent the sum of the three components). The same is plotted in Fig. 5.9 in which now the isolation variable maximum is 0.15. The behaviour of the systematic uncertainty studied is very similar in both cases (isolation variable maximum of 0.1 or 0.15). The systematic uncertainty due to the QCD background subtraction decreases with the p_T threshold. It ranges from a 2% to 0.4% for p_T threshold of 20 GeV and 30 GeV, for an isolation threshold of 0.1, and from 2.8% to 0.6% for an isolation threshold of 0.15. The uncertainty due to the modeling of ISR effect varies from $\approx$ 1% to $\approx$ 2.8%, not depending significantly on the isolation threshold. Finally, the uncertainty due to the PDFs diminishes softly from 1.8% to 1.4% with no dependence on isolation. The total error for the isolation threshold at 0.1 varies from $\approx$ 5% to $\approx$ 4.8% when changing the p_T threshold from 20 GeV to 30 GeV, with a minimum at $\approx$ 4.2% for a p_T threshold of 25 GeV. The corresponding numbers for iso $<$ 0.15 are a variation from 5.8% to 5% with a minimum of 4.4% at 25 GeV.

From these plots we can conclude that we get a minimum of the systematic uncertainty when the p_T requirement is set to approximately 25 GeV. We also observe a reduction in the systematic uncertainties when the isolation cut is set to 0.1.

The selection process to get an optimal sample of W candidates to compute the cross section production can be summarized as follows:

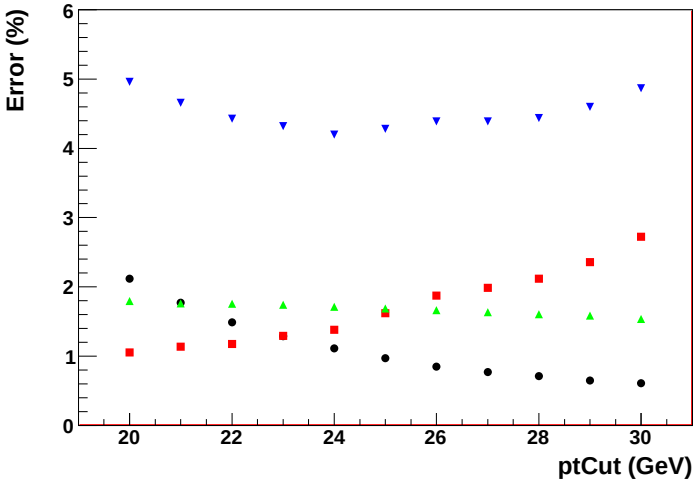

Figure 5.8: - Variation of the systematic uncertainty studied in this section with the p_T requirement applied. Black dots are used for the QCD background modeling systematic uncertainty, the red squares for the ISR effect one, green triangles for the PDF one and blue triangles to represent the sum of the three components. An isolation variable maximum of 0.1 is applied in this plot.

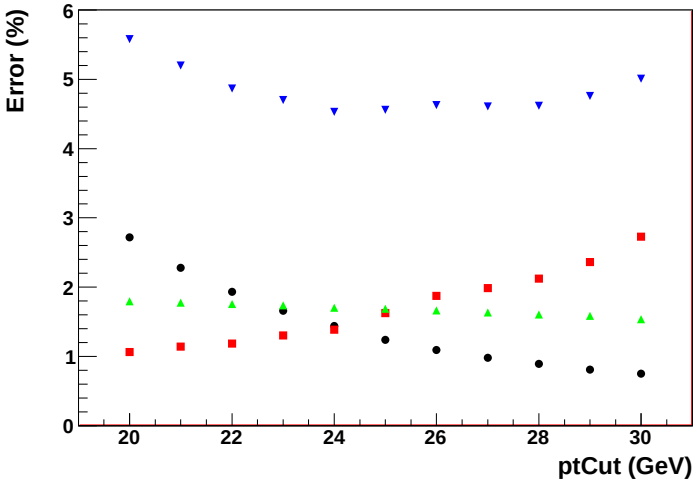

Figure 5.9: - Variation of the systematic uncertainty studied in this section with the p_T requirement applied. Black dots are used for the QCD background modeling systematic uncertainty, the red squares for the ISR effect one, green triangles for the PDF one and blue triangles to represent the sum of the three components. An isolation variable maximum of 0.15 is applied in this plot.

1. The event has to be selected online by any of the HLT triggers seeded by a muon identified in the muon spectrometer (HLT_Mu9, HLT_Mu11 or HLT_Mu15).
2. We define a W candidates as an event with a reconstructed muon.
3. Contribution from Drell-Yan events is suppressed by removing events with two high p_T muons (20 GeV as threshold for the highest p_T muon and 10 GeV for the other one).
4. The reconstructed muon has to fulfill the quality criteria presented before (muon reconstruction and identification requirements).
5. Pseudorapidity of the muon should be below 2.1, to match the region where the muon trigger is more stable.
6. Finally, the selection requirements studied in this section are applied. Muon p_T should be bigger than 25 GeV and the isolation variable should be below 0.1.

This selection procedure is applied to the full 2010 data sample.

The breakdown of the data reduction at the different stages of the selection is summarized in Table 5.3 both for the total sample of muon events, and splitted by the muon charge.

Event Sample	Events with $\mu^{\pm}$	Events with μ^{+}	Events with μ^{-}
Candidates	3490239	1812890	1677349
Triggered	3174722	1644727	1529995
DY Rejection	3133420	1623994	1509425
Muon ID	2618199	1346297	1271902
$\|\eta\| < 2.1$	2527047	1298777	1228270
$p_T > 25$ GeV	412266	222080	190186
$I^{\text{rel}}_{\text{comb}} < 0.10$	166457	97533	68924

Table 5.3: Data reduction at every step of the selection process. The number of events is given for the whole muon data sample, and also separated by the muon charge.

After the selection process just described, 166 457 events are selected, 97 533 of them with a positive charged muon and 68 924 with a negative charged muon. The estimated number of events of the different processes (signal and background) to the $W \to \mu\nu$ candidate sample are summarized in Tab. 5.4.

source	$N_{\rm bg}$ in 36.1 pb^{-1}	$N_{\rm bg}/(N_W + N_{\rm bg})$	$N_{\rm bg}/(N_W + N_{\rm bg})$ (MET> 20 GeV)
QCD multi-jet	8831	5.1%	0.6%
$Z \to \mu\mu$	6130	3.5%	2.9%
$W \to \tau\nu$	4623	2.7%	2.9%
$Z \to \tau\tau$	910	0.5%	0.2%
WW+WZ+ZZ	205	0.1%	0.1%
$t\bar{t}$	592	0.3%	0.4%
EWK + $t\bar{t}$	12538	7.1%	6.4%
total	21434	12.2%	7.0%
$W \to \mu\nu$ signal	152676	87.8%	93.0%

Table 5.4: Estimates of signal and backgrounds in the $W \to \mu\nu$ candidate sample, based on Monte Carlo simulations. The second column gives the number of events expected from each type of processes in a 36 pb^{-1} sample, based on the theoretical predicted cross sections. The third column shows the fraction of each of the processes that contributes to the candidate sample. The last column indicates the fraction of events in a reduced phase space MET > 20 GeV.

5.4.4 Acceptance and efficiencies computation

In this section we will compute the detector acceptance (A), fraction of the W phase space accessible by the CMS detector, and the efficiency of the event selection (ϵ) for W events within the detector acceptance.

The detector acceptance has to be evaluated necessarily by MC simulations. The detector efficiency can be evaluated either by MC simulation or with control samples. The approach we have followed in this thesis is to compute by MC a single factor, A_w, accounting for the two effects. This A_w factor will be in fact equivalent to the product $A \times \epsilon$. Any deviation of the MC description of the detector efficiency will be later evaluated with independent control samples and appropriate correction terms are derived to modify the A_w factor just computed.

However it is difficult to have a perfect description of the MET behaviour in the MC. Thus, remaining discrepancies between data and MC (or remaining miscalibrations in the MC) of the MET scale and resolution will be corrected with information from $Z \to \mu\mu$ samples.

The signal acceptance times efficiency (A_w) term is computed with the $W \to \mu\nu$

Acceptance × efficiency (A_w)		
$W^{\pm}$	W^{+}	W^{-}
0.4071 ± 0.0003	0.4134 ± 0.0004	0.4021 ± 0.0004

Table 5.5: W signal acceptance and efficiency (A_w) factors evaluated on the MC POWHEG sample. The errors shown are the statistical error due to size of the MC sample.

POWHEG reference sample listed in Tab. 5.2. An event lies within the detector acceptance and is selected if the pseudorapidity of the reconstructed muon is below 2.1 (in absolute value) and its transverse momentum is higher than 25 GeV and passes all the selection requirements. This factor is computed as

$$A_w = \frac{N_{sel}}{N_{gen}}.$$

The values of this factor are shown in Tab. 5.5.

Single muon efficiencies prediction from the MC is probed in $Z \to \mu\mu$ samples. Muon efficiencies are evaluated with a Tag and Probe method both in a $Z \to \mu\mu$ MC sample and in data, and a ρ factor is defined as the efficiency in data over the efficiency in MC. This ρ factor will be used to correct the MC prediction of the A_w term. When doing this we extract the efficiencies from a sample of muons kinematically very similar to those from the W decay and exploits the relatively clear selection of Z candidates. The possible presence of background processes in the selected event sample is also taken into account. The method developed is explained in detail in (77, 78). ρ factors are computed for different regions in pseudorapidity to reproduce the real structure of the detector (lower efficiency in the transition regions between the barrel wheels and the barrel and endcap regions) and split by charge to account for possible asymmetries in the detection (selection of positive and negative muons). The overall ρ factors are shown in Tab. 5.6.

A small fraction of muon events are lost because of L1 muon trigger prefiring, i.e., the assignment of a muon segment to an incorrect bunch crossing, occurring with a probability of a few per mille per segment. The effect is only sizable in the drift-tube system. Since this effect is not accounted for in the efficiency from tag and probe we have to consider it separately (we correct the ρ factors by a factor 0.995).

	μ^+	μ^-	$\mu^\pm$
ϵ_{data}	$(85.98 \pm 0.38 \pm 0.72)\%$	$(85.00 \pm 0.36 \pm 0.72)\%$	$(85.48 \pm 0.27 \pm 0.72)\%$
ϵ_{MC}	$(89.25 \pm 0.05)\%$	$(89.38 \pm 0.05)\%$	$(89.32 \pm 0.04)\%$
ρ	$(96.33 \pm 0.43 \pm 0.81)\%$	$(95.09 \pm 0.40 \pm 0.81)\%$	$(95.70 \pm 0.30 \pm 0.81)\%$

Table 5.6: Efficiencies and corrections factors for positive, negative and full sample of muons. The first uncertainty shown is statistical and the second systematical.

Acceptance		
$W^\pm$	W^+	W^-
0.4638 ± 0.0003	0.4706 ± 0.0004	0.4570 ± 0.0004

Table 5.7: .

To have an idea of how much of the A_w term is due to detector acceptance and to detector efficiency, we include also the values of the acceptance. We will define that an event lies in the detector acceptance if it has a muon with a $p_T > 25$ GeV in the region $|\eta| < 2.1$, at generator level:

$$A = \frac{N_{gen}(p_T^{gen} > 25\ GeV, |\eta| < 2.1)}{N_{gen}}.$$

The values of the acceptance are given Tab 5.7.

5.4.5 W signal extraction

We now focus on the numerator of the previous formula to compute the production cross section:

$$\sigma = \frac{(N_{signal} - N_{background})}{A\epsilon\mathcal{L}}. \tag{5.9}$$

To evaluate the number of signal and background events in the selected sample, we focus on a highly discriminant variable, the missing transverse energy (MET). This variable allows to clearly discriminate signal from the dominant background QCD. The strategy will be to perform a binned likelihood fit to the observed MET distribution of a sum of the different contributions already discussed (signal, QCD background, EWK backgrounds, dibosons and $t\bar{t}$). The shape of the signal and background MET distribution are modeled as individual templates and the amount of different contributions

is allowed to float in the fit. The way the MET templates are built for the different contributions and the details in the fit will be shown in the following subsections.

MET signal template. MET template for signal can be drawn from MC simulation. ElectroWeak processes are well known, and the description of the MET shape for the signal contribution can be safely extracted from the MC prediction. To improve the MET description, any possible miscalibration of the E_T scale and resolution in the MC, it is corrected for with information from $Z \to \mu\mu$ event (79).

W and Z bosons have similar recoil models, they are produced at a similar Q^2 and their production and decay processes are similar. The idea is to extract information from $Z \to \mu\mu$ events where the whole event can be reconstructed in the transverse plane and use it to model the behaviour of W events. The recoil to the Z, in the transverse plane, i.e. all the energy of the event but the two muons, can be decomposed in two orthogonal directions: parallel and perpendicular to the Z direction. The parallel component of the recoil is due to initial state gluons radiated from the quarks that produce the Z. This emission balances the p_T of the boson, and therefore the average value of $\text{MET}_{\parallel}$ is expected to increase with Z-p_T. The second component ($\text{MET}_{\perp}$), is due to multiple interactions and remnants of the beam particles involved in the Z production. The average value of $\text{MET}_{\perp}$ is essentially zero with a certain resolution effect.

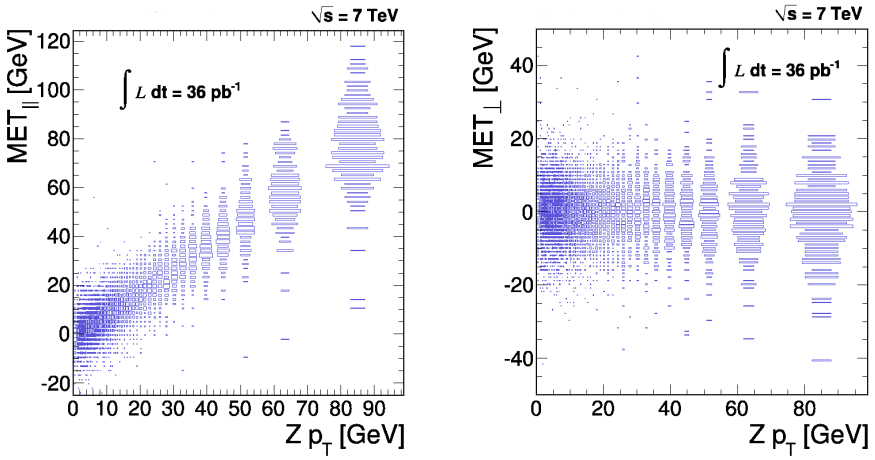

Figure 5.10: Recoil distribution in $Z \to \mu\mu$ events: the plots show the parallel (left) and perpendicular (right) components, as a function of the p_T of the boson.

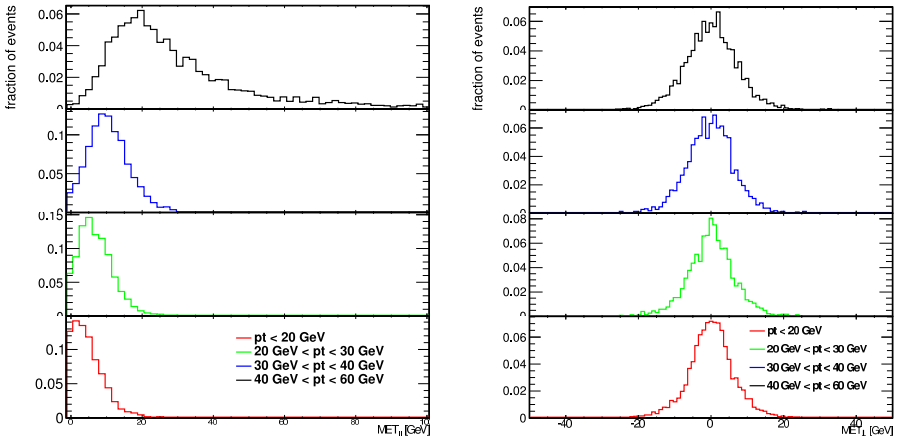

Figure 5.11: MET distribution in Z events: Parallel (left) and perpendicular (right) components, plotted for several intervals of the boson p_T (from bottom to top), $[0-20]$ GeV, $[20-30]$ GeV, $[30-40]$ GeV and $[40-60]$ GeV.

In Fig. 5.10 the parallel and perpendicular components of the Z boson recoil as a function of the Z p_T in real data is shown. The information is stored binned in the Z boson p_T, with a variable binning to have a uniform population in every bin. For every MC simulated W with a given p_T, we sort randomly a value for the $\text{MET}_{\parallel}$ from the slice corresponding to the same p_T of the Z boson. The perpendicular component, since it does not depend on the boson p_T, is sorted from the global distribution. In Fig. 5.11 the $\text{MET}_{\parallel}$ and $\text{MET}_{\perp}$ projections for several bins in the Z boson p_T are shown. The MET so constructed can be combined with the rest of the information of the simulated event (reconstructed muon, angle between the muon and the MET, etc.) to calculate any derived quantity. The resulting W template is presented in Fig. 5.12 compared to the prediction of the MC simulation. One can appreciate a slight broadening of the shape towards higher values of MET.

QCD MET template

It is known that MC simulation for QCD related processes is not to be accurate enough as to fully rely on them to draw conclusions about their characteristics. We will define an independent control sample from which this shape can be extracted. The distribution of the isolation variable is shown in Fig. 5.13 for the MC simulation. The threshold in the isolation variable for the signal selection was set at 0.1. One can appreciate that events with isolation variable bigger than 0.2 are mostly QCD events.

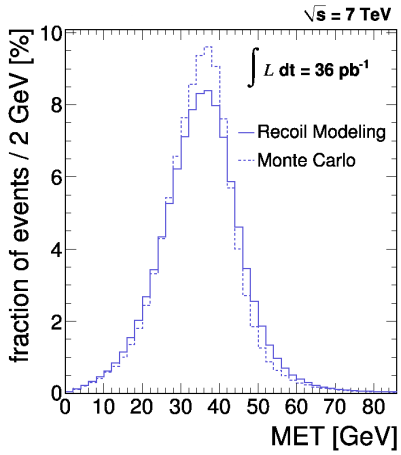

Figure 5.12: The plain prediction from the MC (dashed-line) is compared with the template incorporating the Z recoil correction (solid line).

We can build a MET template for QCD background as the MET shape of the non-isolated events. However, as shown in Fig. 5.14, the non-isolated QCD events do not behave like the isolated QCD events. In particular, for the simulation, MET varies with the isolation variable. This correlation can be removed applying a correction to MET which varies linearly with I_{rel}^{comb}. Under these conditions a new corrected template is obtained applying this correction to the events in the non-isolated region. The physical reason of this correlation is the dependence of the isolation quantity (energy in the calorimeters and p_T of the tracks) on the total activity of the event (ΣE_T). Fig. 5.15 shows the corrected template that is now much closer to that shape of the isolated region. The same kind of behaviour is observed in the data (Fig. 5.14). There is also a linear dependence of the MET variable with the isolation as in MC, although with a slightly different slope. Following the same spirit that in MC we will extract in data the MET shape from the non isolated region, once the correlation is corrected. The MET-isolation dependence was obtained as $\mathrm{MET}' = \mathrm{MET}(1 + \alpha \mathrm{I}_{\mathrm{rel}}^{\mathrm{comb}})$ with $\alpha \approx 0.19$. The QCD templates obtained using this procedure are shown in Fig. 5.16. There it is plotted the MET shape from the non isolated region (dashed line) and how it is modified (dots) once the MET-iso correlation is removed. This will be the final template used in the W cross section measurement. It is also compared with the MC prediction for the isolated region. They do not fully agree, but, as it was already discussed, it is not expected that QCD MC reproduces completely the behaviour of the data. Finally, the

red rectangles indicate the systematic uncertainty assigned to the QCD MET shape, evaluated as discussed in Section 5.5.

QCD template is assumed to be the same in shape and magnitude for both charges. This assumption is shown to be correct within the errors associated to the QCD shape.

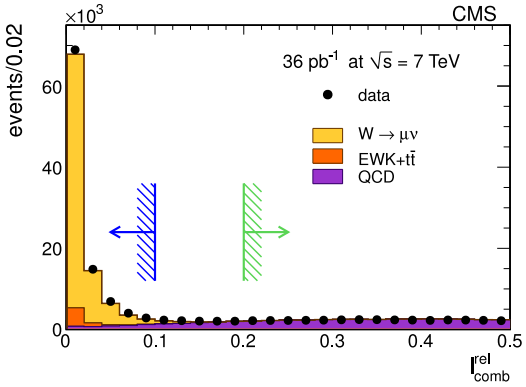

Figure 5.13: Isolation variable distribution. - Isolation distribution of candidates with a good quality muon of $p_T > 25$ GeV in the fiducial region $|\eta| < 2.1$. Dots represent the data and the solid histograms the contribution from the different SM processes, evaluated by MC and normalized to the theoretical cross sections. The blue arrow ($I_{rel}^{comb} < 0.10$) indicates the threshold that defines the signal region. The green arrow defines a background control region, $I_{rel}^{comb} > 0.20$.

MET Template for the other ElectroWeak processes. We rely on the MC prediction to derive the shape of the MET distribution of the candidates originating in the other electroweak processes ($W \to \tau\nu$, $Z \to \tau\tau$ and $Z \to \mu\mu$). A similar procedure as the one followed with the signal can be envisaged. However the improvement in the MET description in this case is found to be negligible and within the errors associated to this modeling, with a negligible impact in the final result (the electroweak background is of the order of 5% in the selected sample). The template used for the other electroweak processes is shown in Fig. 5.17. The absolute level of background will be derived from the fit.

MET templates for $t\bar{t}$ and diboson events. Contribution from dibosons production and $t\bar{t}$ production to the final sample of candidates is negligible. The shape is obtained from MC simulation and their absolute normalization from theoretical predictions.

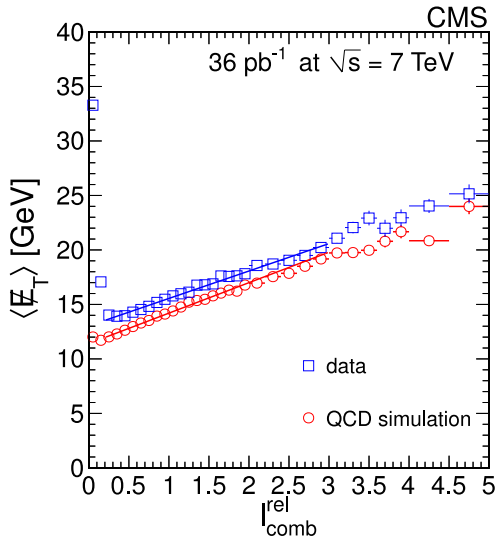

Figure 5.14: - Correlation of MET with the isolation variable. Red circles are the prediction from QCD events and the blue squares the data behaviour. A linear dependence is observed and extracted to improve the template for the QCD sample. The first point (low isolation variable) presents a high MET value. This is due to the presence of signal events, entering specially in this bin.

5.4.6 Cross section measurement

Once we have modeled the behaviour of the MET for the different contributions we can proceed to the W cross section evaluation. The total contribution to the number of selected events is divided in the 4 previously described: signal, QCD, EWK and others. Technically speaking, we perform a binned fit to minimize the likelihood between data and the sum of the contributions. In each bin, the number of expected events is:

$$N(\mathrm{MET}) = \mathrm{N_W}(\mathrm{MET}) + \mathrm{N_{EWK}}(\mathrm{MET}) + \mathrm{N_{QCD}}(\mathrm{MET}) + \mathrm{N_{Others}}(\mathrm{MET}) \qquad (5.10)$$

The W and EWK terms are expressed in terms of their cross sections, their acceptances and efficiencies ($\mathcal{A}_W$ and $\mathcal{A}_{EWK}$ include both factors) and their probability distribution functions (p.d.f) in the MET variable (from the templates), $\mathcal{F}_W$ and $\mathcal{F}_{EWK}$. The EWK contributions, $Z \to \mu\mu$, $Z \to \tau\tau$ and $W \to \tau\nu$ are normalized to the $W \to \mu\nu$ signal, through their theoretical cross section ratio ($\mathcal{K}$ factor). The QCD contribution is described as well in terms of a p.d.f. (template) on MET ($\mathcal{F}_{QCD}(\mathrm{MET})$) and a constant ($\mathcal{N}_{QCD}$) setting the absolute background level. The contribution given by the rest

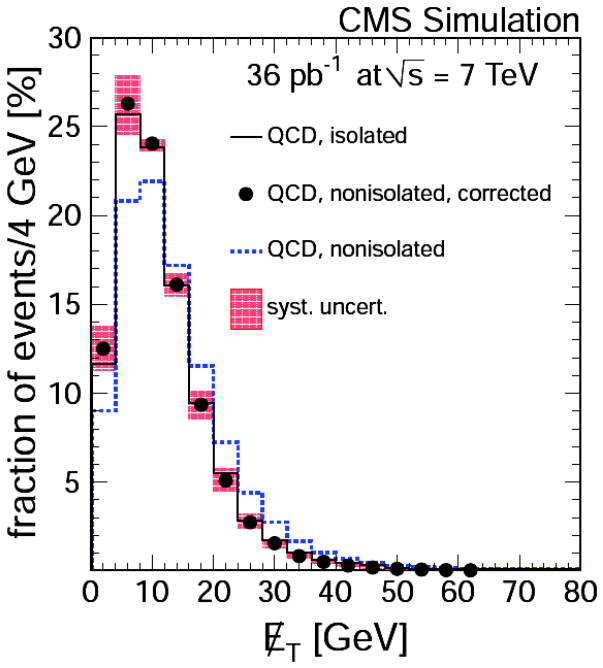

Figure 5.15: Distribution of the corrected MET for selected events with a non isolated muon (black points) superimposed on the distribution of uncorrected MET for the same events (blue, dashed line) and MET for events with an isolated muon (black, solid histogram). All distributions are from simulated QCD events. The shaded area represents the systematic uncertainty due to corrections with factors $\alpha \pm \Delta\alpha$, for $\Delta\alpha = 0.08$.

of the processes, as explained before, is computed from MC and theoretical predictions and is constant through the fit process. We can compute the expected number of events in a certain region of the MET spectrum as:

$$N(\text{MET}) = \{\sigma_{\text{W}} \times [\mathcal{A}_{\text{W}} \cdot \mathcal{F}_{\text{W}}(\text{MET}) + \mathcal{K} \cdot \mathcal{A}_{\mathcal{EWK}} \cdot \mathcal{F}_{\mathcal{EWK}}(\text{MET})] + \mathcal{N}_{\text{QCD}} \cdot \mathcal{F}_{\text{QCD}}(\text{MET})\} \times \mathcal{L}_{int} + \mathcal{N}_{Others} \quad (5.11)$$

This fitting process is performed simultaneously for the positive and negative channels. The fitting function can be expressed in terms of two different sets of parameters:

1. the total W cross section ($\sigma(\text{W}^+) + \sigma(\text{W}^-)$) and the ratio $R = \sigma(\text{W}^+)/\sigma(\text{W}^-)$, together with the overall normalization of QCD events ($\mathcal{N}_{QCD}$),

2. the individual $\sigma(\text{W}^+)$, $\sigma(\text{W}^-)$ and the overall normalization of the background.

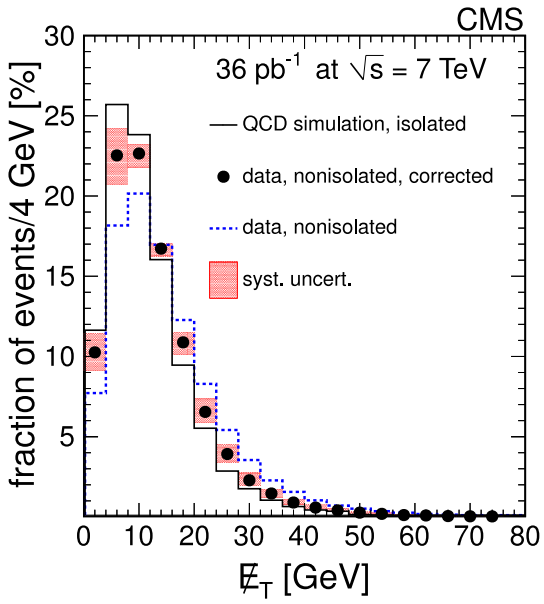

Figure 5.16: Distribution of the corrected MET for selected events with a non isolated muon in data (black points) superimposed on the uncorrected MET distributions for data (blue dashed line) and simulated QCD events (black, solid histogram, same as the black, solid histogram in Fig. 5.15. The shaded area represents the systematic uncertainty due to corrections with factors $\alpha \pm \Delta\alpha$, for $\Delta\alpha = 0.08$.

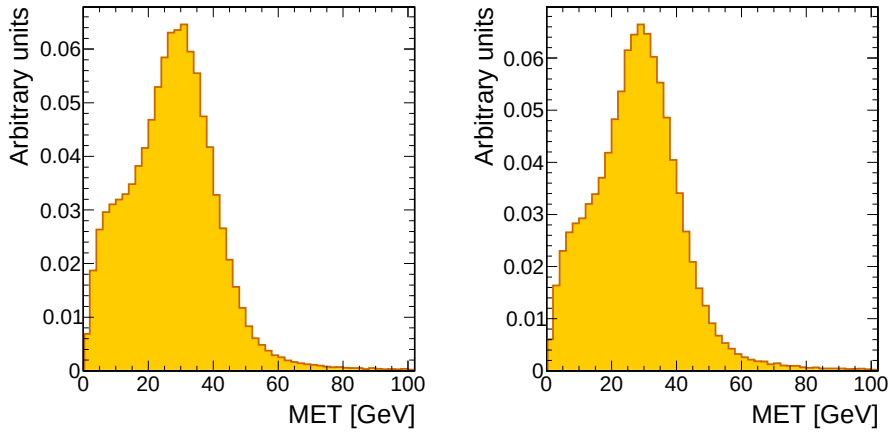

Figure 5.17: Sum of all MET distribution of the different electroweak processes considered (appropriately normalized, using the theoretical cross section for each electroweak process) for minus (left) and plus (right) processes. In both figures the area is normalized to the unity.

The fit is performed over the MET range $[0, 200]$ GeV. Fitted W yields of 141226 ± 376 (total), 84315 ± 290 W^+ and 56911 ± 239 W^-, events are obtained.

The total W cross section, and the individual W^+ and W^- cross sections, as well as their ratio, obtained from the fit are presented in Table 5.8. The errors shown are only statistical.

In Fig. 5.18, the experimental distribution in the MET variable together with the other contributions after the fit is shown, both in linear and logarithmic scale. A good agreement is observed after the fit in all the MET range plotted.

The W-signal extraction procedure can also be formulated in terms of a fit to the M_T distribution, simply deriving the corresponding M_T templates from those in MET (equation 5.5). In Fig. 5.19 the results of the fit in the M_T variable are plotted. Numerical results from the M_T fit are also given in Tab. 5.8. The two sets of results are in full agreement. As we will see as one of the conclusions of this Chapter, a requirement of $M_T > 50$ GeV will guarantee the selection of an almost background free W signal. We can notice in the Fig. 5.19 the decrease of the background when this later requirement is set.

5.5 Systematical uncertainties

The measurement presented is affected by the uncertainties in all the different elements that take part of it.

In this section we will show the estimation of these systematical uncertainties. They are grouped as experimental uncertainties and theoretical ones, depending on the source of the uncertainty. Luminosity uncertainty (dominant in the measurement) is consider separately.

As experimental sources we will consider the uncertainties on the lepton reconstruction and identification, the trigger prefiring, the muon momentum scale and resolution and the signal and background modeling.

The total uncertainty (statistical and systematic) on the correction factors (ρ) is taken as the systematic uncertainty due to muon efficiency (reconstruction, identification, selection, isolation and trigger).

The muon p_T scale and resolution affects the measurement of the W boson cross section production through the selection based on the muon p_T as well as in the MET

Fit in MET distribution	
$\sigma(\mathrm{W})$ (nb)	10.101 ± 0.027
$R = \sigma(\mathrm{W}^+)/\sigma(\mathrm{W}^-)$	1.421 ± 0.008
QCD normalization factor	1.37
$\sigma(\mathrm{W}^+)$ (nb)	5.929 ± 0.021
$\sigma(\mathrm{W}^-)$ (nb)	4.172 ± 0.018
QCD normalization factor	1.37
Fit in M_T distribution	
$\sigma(\mathrm{W})$ (nb)	10.150 ± 0.027
$R = \sigma(\mathrm{W}^+)/\sigma(\mathrm{W}^-)$	1.418 ± 0.008
QCD normalization factor	1.33
$\sigma(\mathrm{W}^+)$ (nb)	5.953 ± 0.020
$\sigma(\mathrm{W}^-)$ (nb)	4.197 ± 0.018
QCD normalization factor	1.33

Table 5.8: Total W, W^+ and W^- production cross section (times the Branching Ratio of the W decaying into a muon and a neutrino) and ratio between W^+ and W^- cross sections. The QCD normalization factor is also presented. The upper set of numbers are obtained from a fit to the MET distribution. The lower one was obtained from a fit to the M_T distribution.

shape. The SIDRA method presented in Chapter 4 is used to get the corresponding scale and resolution factors binned in pseudorapidity and azimuthal angle. These factors are applied to the muons coming from the MC simulation, obtaining a new p_T, and thus new templates and efficiencies. In these conditions, we repeat the cross section fit previously exposed. The error associated to this source is computed as the difference of this value with the reference one. A 0.3% variation is obtained.

The mismodeling of the MET shape for the signal is another source of uncertainty, related with the MET scale and resolution. It is estimated as the difference in the cross section measurement using the reference template for the signal and the one provided by MC. This difference is shown to be of 0.2%.

The systematic uncertainty in the background subtraction takes into account the effect of possible mismodeling of the MET shape of the QCD component on the cross section measurement. MC studies indicated that the optimal value of α is 0.24 to correct for the isolation-MET correlation. Fig. 5.20 shows the MC prediction for the

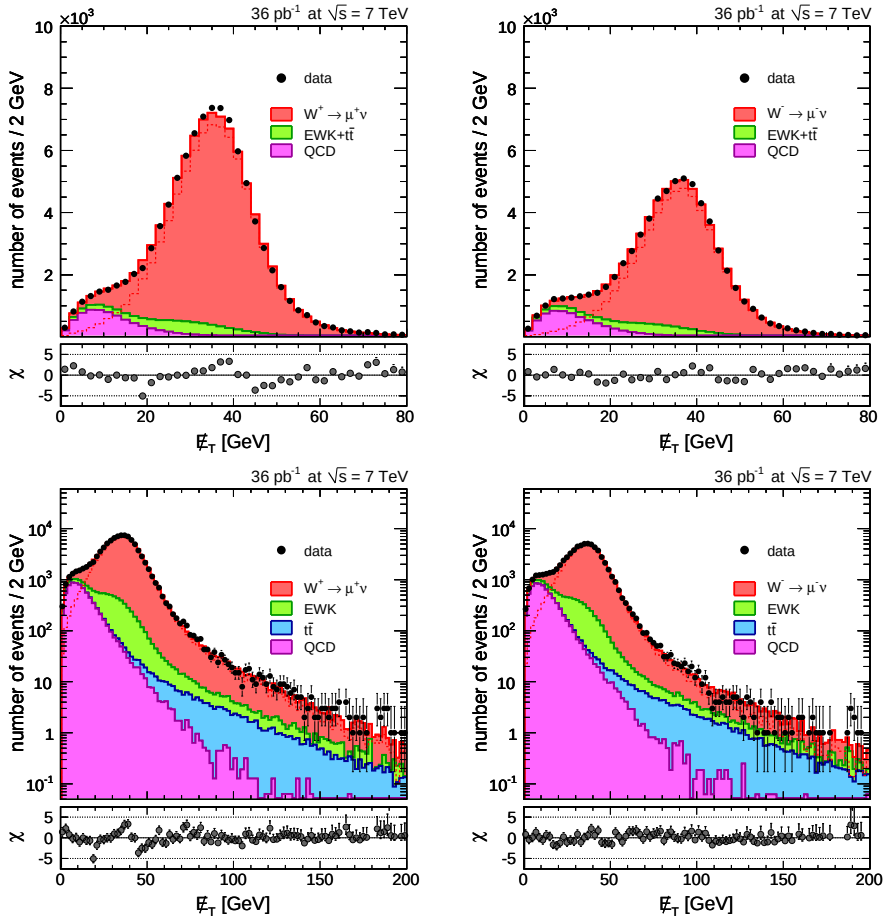

Figure 5.18: W^+ (left plots) and W^- (right plots) experimental distributions (black dots) of MET, both in linear and logarithmic scale, together with the fitted contributions from the different processes (shown stacked): W signal (red histogram), other EWK processes (green histogram), QCD background (pink histogram) and $t\bar{t}$ (blue histogram).

isolated QCD events (black solid line), the prediction derived from the non-isolated region (red dashed line) and the prediction once the correlation has been corrected (green line). The prediction in this last case is much closer to the prediction from the isolated part. Predictions using other values of α (from 0.16 to 0.32) are also shown in the Figure. A variation of $\Delta\alpha = \pm 0.08$ successfully covers the MC prediction for the isolated region over all MET interval. To evaluate the uncertainty associated to the method we fit the MET spectrum with two extreme MET shapes using modified α correction factors according to this variation.

We consider a 0.5% to accout for the L1 muon trigger prefiring, i.e., the assignment

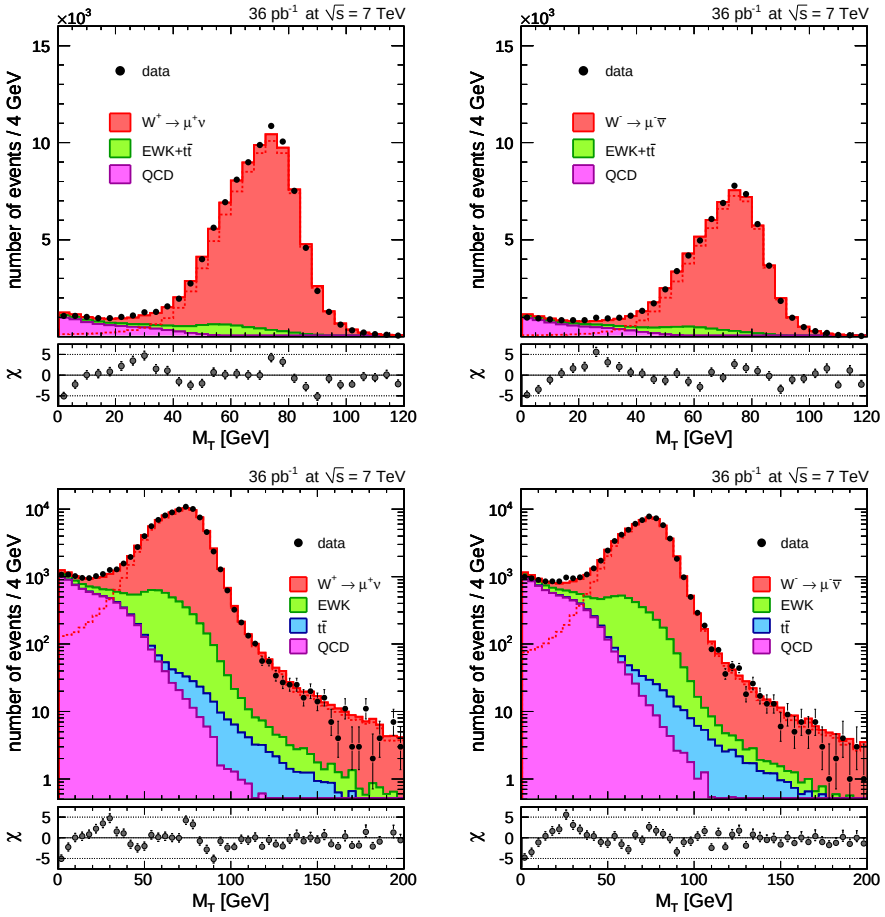

Figure 5.19: W^+ (left plots) and W^- (right plots) experimental distributions (black dots) of M_T, both in linear and logarithmic scale, together with the fitted contributions from the different processes (shown stacked): W signal (red histogram), other EWK processes (green histogram), QCD background (pink histogram) and $t\bar{t}$ (blue histogram).

of a muon segment to an incorrect bunch crossing. Since this effect is not consider in the Tag and Probe method, the uncertainty due to this effect should be accounted for separately. A 0.5% uncertainty is estimated for this effect.

As theoretical uncertainties we consider the PDF uncertainty and other theoretical effects such as ISR (initial state radiation) and FSR (final state radiation) effects. The MC generator used for the signal simulation is POWHEG interfaced with PYTHIA for parton showering and hadronization and fragmentation. PYTHIA is "tuned" to better describe the data by modifying a hundred of parameters related to not well know processes in the collisions, such us multiple interactions, the underlying event,

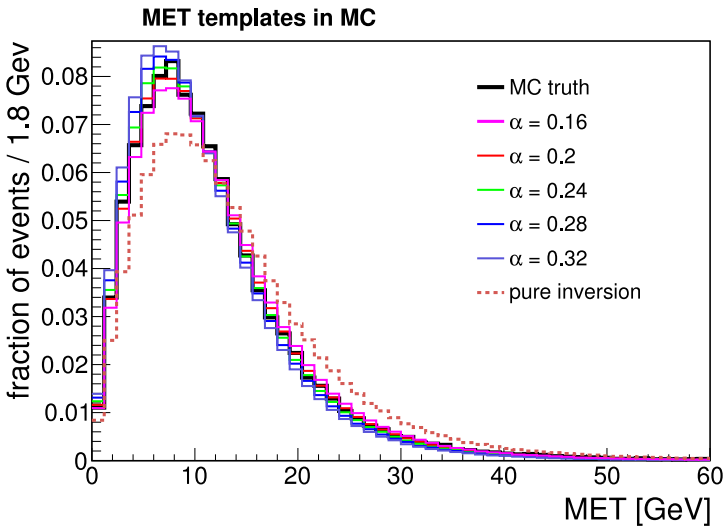

Figure 5.20: MET template for the QCD MC sample (black). With other colors, variations of this prediction obtained with different α values. The optimal value for α in MC is computed to be 0.24. In this plot we observe how a variation of 0.08 around the α central value completely covers the MC prediction. The MC truth refers to the isolated MC.

etc. Several parameters control the ISR and FSR effects. Variation of standard tune of PYTHIA in these parameters with respect to the official CMS tune is the way chosen to estimate the uncertainty given by these effects.

The impact of these in the total W cross section and on the individual W^+ and W^- as well as on the ratio is presented in Tab. 5.9.

The experimental error is of the order of 1.1% (total W cross section) and the theoretical one 1.1%. The total error excluding luminosity is 1.6%. Individual cross section theoretical uncertainties are slightly larger (the experimental are essentially the same). Ratios are more affected by theoretical uncertainties compared to the total cross sections. A 4% uncertainty in the recorded luminosity is assigned.

As a comparison, the 3.1% uncertainty of the previous measurement is now reduced to 1.6%, thanks to the use of new techniques that profit from the big amount of data available (ten times bigger than in the previous measurement of the W boson cross section by CMS).

Source	W	W^+	W^-	W^+/W^-
Lepton reconstruction & identification	0.9	0.9	0.9	1.3
Trigger prefiring	0.5	0.5	0.5	0
Momentum scale & resolution	0.3	0.3	0.3	0
MET scale & resolution	0.2	0.2	0.1	0.1
Background subtraction / modeling	0.4	0.4	0.5	0.1
Total experimental	1.1	1.1	1.2	1.3
PDF uncertainty for acceptance	0.8	0.9	1.5	1.9
Other theoretical uncertainties	0.8	0.9	0.8	0.8
Total theoretical	1.1	1.3	1.7	2.1
Total (excluding luminosity)	1.6	1.7	2.1	2.5

Table 5.9: Systematic uncertainties in percent for individual W cross sections and ratios in the muon channel. A common luminosity uncertainty of 4% applies to all cross sections.

5.6 Kinematic distributions

In this Section we show several kinematic distributions, necessary to completely characterize the $W \to \mu\nu$ events produced in the collisions. All the distributions are normalized to the output of the fit.

In Fig. 5.21 the p_T distribution of the W candidates is shown. The 25 GeV selection requirement p_T can be noticed here. The agreement data-MC in this plot is not perfect, with visible disagreements in the whole range. However, the distribution for the "golden candidates" (if we additionally require the transverse mass of the candidate to be higher than 50 GeV) shows a good agreement with MC. In Fig. 5.22 we show the p_T distribution of this subsample.

In Fig. 5.23 (left), the pseudorapidity distribution of the W candidates is presented. It is possible to notice the efficiency drop in the inter-barrel regions and the overlaps. Fig. 5.23 (right) shows the azimuthal angle distribution for all the W candidates. This last distribution is flat in this angle, showing that there is not a preferred decay directionality in the transverse plane of the W bosons in CMS.

The acoplanarity variable measures the angle between the muon and the MET of the event. It was defined as (5.6). In W events the muon and the neutrino are usually back-to-back (differ by π radians in ϕ). As a consequence the acoplanarity of the muon and the MET for these events will be low. The acoplanarity distribution for the W

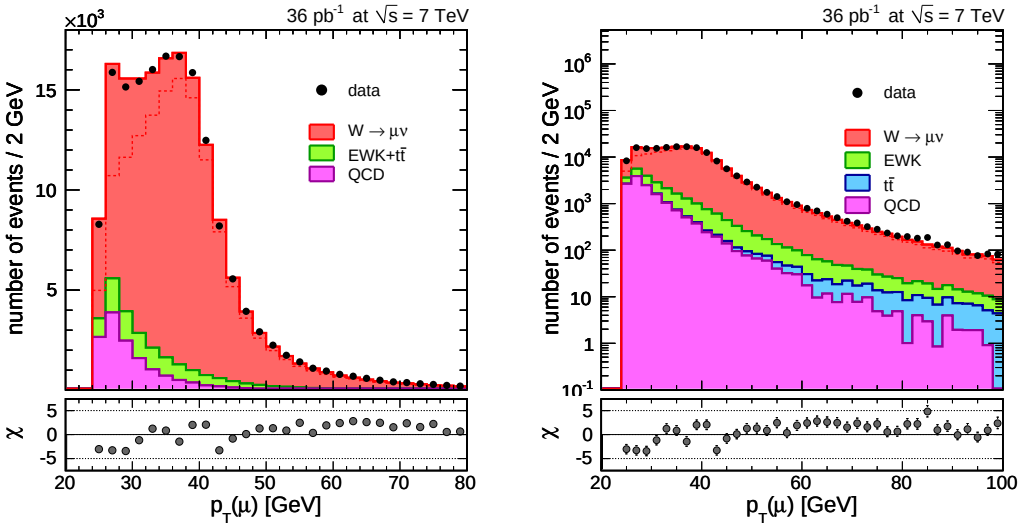

Figure 5.21: Muon p_T distribution for the $W_{\mu\nu}$ sample.

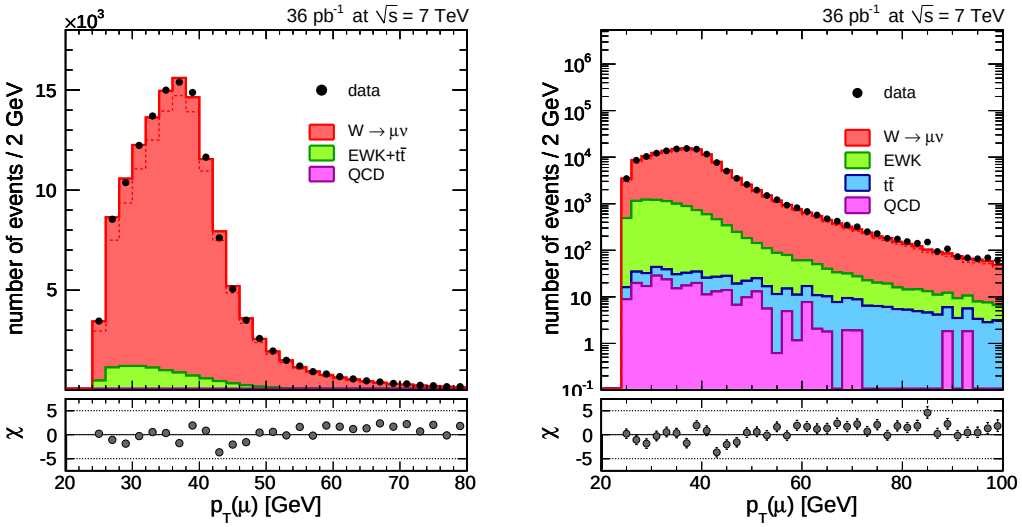

Figure 5.22: Muon p_T distribution for the $W_{\mu\nu}$ sample. Only $W_{\mu\nu}$ candidates with a $M_T > 50$ GeV are included in the plot.

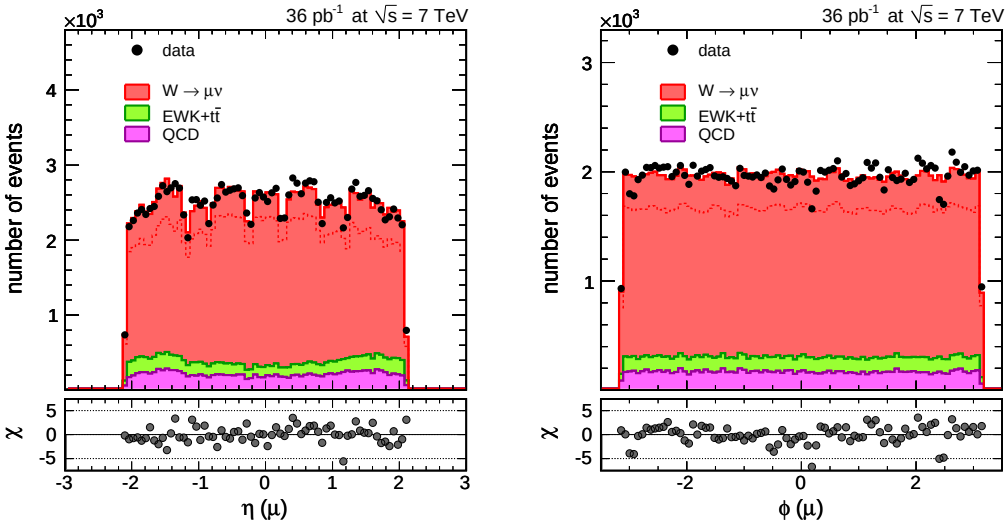

Figure 5.23: Muon η and ϕ distributions for the $W_{\mu\nu}$ sample.

candidates is shown if Fig. 5.24. The signal peaks at low acoplanarity (events are back-to-back) as expected.

In Fig. 5.25 the p_T distribution of the W candidate is shown. In Fig. 5.26 the same variable is plotted for the W golden candidates (those with M_T higher than 50 GeV). A general disagreement is observed in these plot, specially in the low p_T region. Low p_T region is dominated by non-perturbative QCD effects and MC tuning.

The results are summarized in Fig. 5.27 where the total cross section, individual cross sections and ratio computed in this thesis are compared with the theoretical ones and with the published results from CMS.

5.7 Discussion of the results

The results on the W boson production cross section in the muonic channel obtained in this analysis (10.1 ± 0.03 $(stat.) \pm 0.16$ $(sys.) \pm 0.4$ $(lumi.)$ nb) are in agreement with the SM (FEWZ and the MSTW 2008 PDF predict 10.44 ± 0.27 nb). The charged cross sections given in this study (5.93 ± 0.02 $(stat.) \pm 0.1$ $(sys.) \pm 0.24$ $(lumi.)$ nb for the positive and 4.17 ± 0.02 $(stat.) \pm 0.09$ $(sys.) \pm 0.17$ $(lumi.)$ nb) are also in agreement with NNLO SM prediction (6.15 ± 0.17 nb and $4.29 \pm 0.11 \pm$ nb for the positive and negative production respectivily), as well as the ratio measured $1.421 \pm$

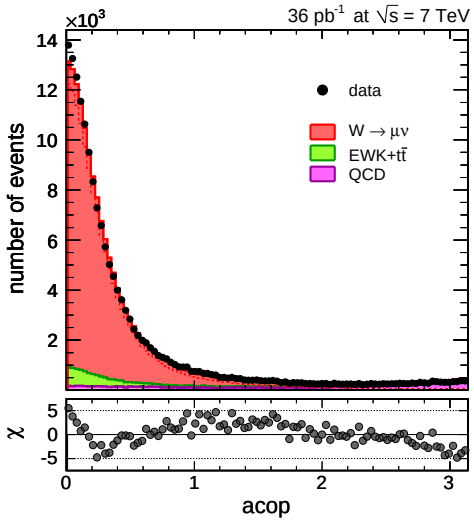

Figure 5.24: Acoplanarity between the muon and the MET distribution for the $W_{\mu\nu}$ sample.

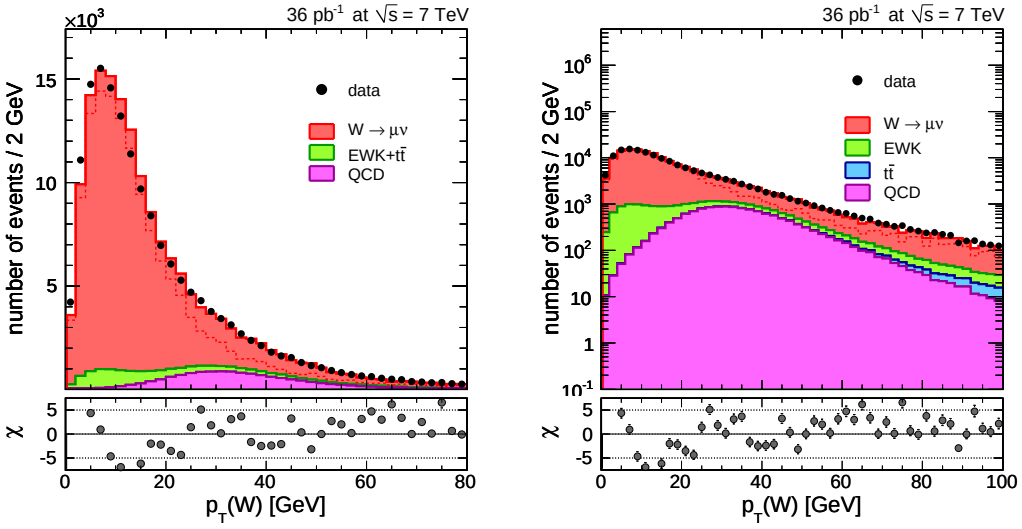

Figure 5.25: Reconstructed W-p_T distribution for the $W_{\mu\nu}$ sample.

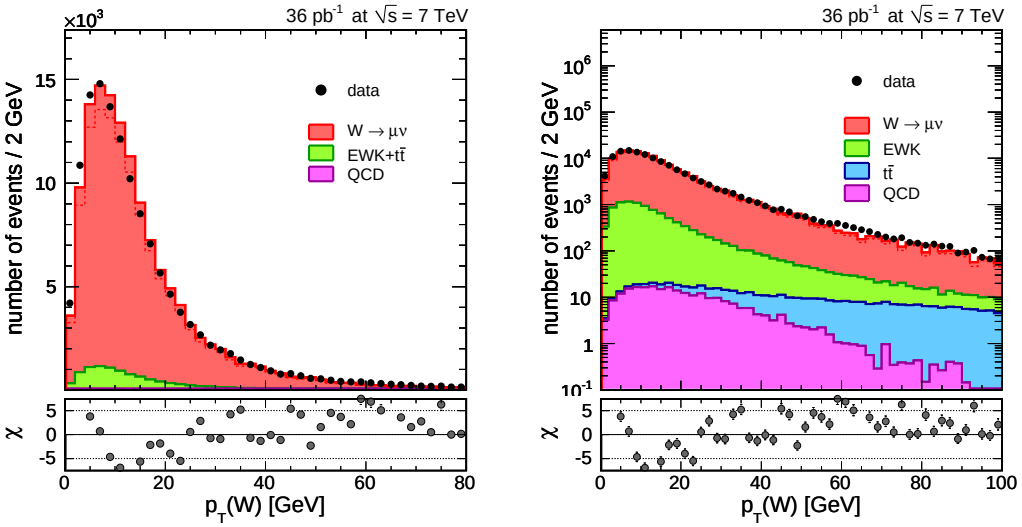

Figure 5.26: Reconstructed W-p_T distribution for the $W_{\mu\nu}$ sample. Only $W_{\mu\nu}$ candidates with a $M_T > 50$ GeV are included in the plot.

$0.008\ (stat.) \pm 0.04\ (sys.)$ (to be compared with 1.43 ± 0.01). The value quoted by the CMS official measurement $10.18 \pm 0.03\ (stat.) \pm 0.16\ (sys.) \pm 0.4\ (lumi.)$ nb is compatible with the one presented in this analysis. The individual cross sections given in this paper $5.98 \pm 0.02\ (stat.) \pm 0.11\ (sys.) \pm 0.24\ (lumi.)$ for the positive and $4.20 \pm 0.02\ (stat.) \pm 0.09\ (sys.) \pm 0.17\ (lumi.)$ for the negative, as well as the ratio positive-negative $1.423 \pm 0.008\ (stat.) \pm 0.04\ (sys)$ are compatible with the results presented in this thesis.

The method proposed for this measurement is similar to the official CMS measurement (80). The only difference between them is in the way the MET template for the signal contribution is built. In the present analysis, a "sampling" of the Z recoil is used to improve the signal MET shape, as previously described. The official CMS measurement also uses the Z recoil to improve the signal MET description, but in this case, instead of randomly select a value of the parallel and perpendicular components following a certain p.d.f., the distributions of the recoil components (parallel and perpendicular to the boson p_T direction) are fitted with a double Gaussian. The mean and width of the Gaussian vary with the boson transverse momentum. For each sample, polynomials are fitted to the extracted mean and width of the recoil distributions as functions of the boson transverse momentum. The ratios of data to simulation fit-

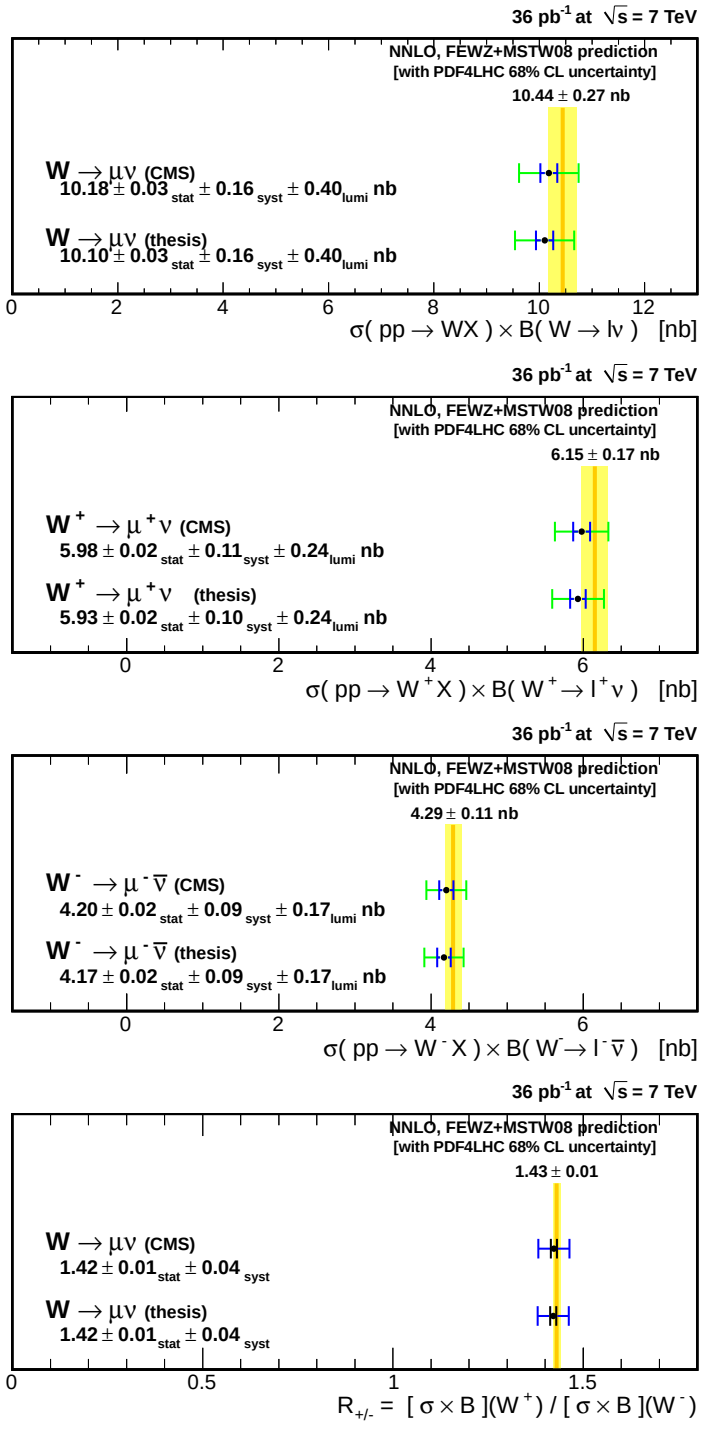

Figure 5.27: Summary of the results shown in this Chapter. Total cross section (upper plot), positive and negative W boson production cross section (medium plots) and the positive-negative ratio (lower plot) are compared with the results from the official CMS measurement and the theoretical prediction (yellow band).

parameters from the Z samples are used as scale factors to correct the polynomials parameters of the W simulated recoil curves. For each W simulated event, the recoil is replaced with a value drawn from the distribution obtained with the corrected parameters corresponding to the W p_T. The MET value is calculated by adding back the energy of the W lepton.

In this analysis we preferred to sample the real Z distribution instead of parametrizing and fitting it so that all effects not accounted for in the Gaussian description are included.

CMS published a first measurement with the first 3 pb^{-1} (57). The method in that analysis to compute the cross section is similar to the one presented in this Chapter, with several differences. The selection requirement were not the same. In order to maximize the number of final events in this measurement it was preferred to set looser requirements on the muon p_T and on the isolation variable (20 GeV and 0.15 respectively). Efficiencies were also computed with the tag and probe method. However, due to the limited statistics available, it was not possible to split by charge, nor to compute it for different regions in pseudorapidity. As a consequence we expect a reduction of the systematic uncertainties in this analysis with respect to the previous one in the background modeling (we are using an optimal requirement based on this uncertainty), and in the muon reconstruction and identification. The muon reconstruction and identification is reduced from 1.5% to 0.9%. The biggest improvement observed is in systematic uncertainty due to background substraction, that amounts now to 0.4%. The good level of understanding of the data also allowed to drop any remaining dependence on the QCD MC for the evaluation of the associated systematic uncertainty that are now evaluated in a more realistic way, based only on data information. In the previous measurement a 2% systematic uncertainty was assigned for the backgound modelling error. A better measurement of the delivered luminosity by the experiment made possible to reduce the 11% luminosity uncertainty of the previous measurement to the 4% that is considered in this measurement. The measurement published with 3 pb^{-1} quoted a value for the cross section of 9.92 ± 0.09 $(stat.) \pm 0.31$ $(syst.) \pm 1.1$ $(lumi.)$ nb, individual cross sections of 5.84 ± 0.07 $(stat.) \pm 0.18$ $(syst.) \pm 0.64$ $(lumi.)$ for the positive and 4.08 ± 0.06 $(stat.) \pm 0.15$ $(syst.) \pm 0.45$ $(lumi.)$ for the negative. The ratio quoted is 1.433 ± 0.026 $(stat.) \pm 0.054$ $(syst.)$. All these results are in agreement with the updated study.

The other general purpose LHC experiment (ATLAS) also published the W boson inclusive cross section in the muonic channel with 36 pb^{-1} (81). Their measurement was in several aspects different to the one presented in this Chapter. The selection process included a requirement on the M_T of 40 GeV. The MC generators used by ATLAS (MCatNLO for signal and $Z \rightarrow \mu\mu$, PYTHIA for other electroweak processes) differ from the ones used by CMS. ATLAS measurement relies on the MC predictions (does not improve the MET description of the signal) and does not perform a fit to extract the normalization of the different contributions (a cut and count method is applied). The results published in this measurement (10.21 ± 0.03 $(stat.) \pm 0.23$ $(syst.) \pm 0.35$ $(lumi.)$ nb) show a bigger uncertainty, with a leading systematic uncertainty of 1.5% from the acceptance computation.

The measurement performed in this analysis shows a result with a precision at a really challenging level of $\approx 2\%$ as a systematic error. With the data collected by CMS (5 fb^{-1}) at the end of 2011 this measurement could be updated. However the limiting luminosity uncertainty and the increasing prescale triggers make this measurement not to be atractive from the physical point of view with 2011 data. With the LHC probably running at 8 TeV in 2012, a new point in the collision energy-cross section could be provided. A better description of the W boson transverse momentum and M_T shape are some of the goals for 2012 related to the W boson production in the collisions.

6

Measurement of associated production of W bosons with charmed jets

At tree level W bosons are produced via quark currents ($q\bar{q}$) . The inclusive production was the subject of the previous Chapter. With increasing recorded luminosity in the experiment, the amount of W bosons produced in association with a jet (group of tracks in the same region of the detector produced by the fragmentation and hadronization of quarks) in the final state starts to be relevant. The study of the associated production of W bosons with charmed jets is the subject of this Chapter.

The associated production of W bosons with charmed jets at LHC differs from other previous experiments due to the high energy of the collisions and the nature of the colliding particles. At very high energy, the presence of sea quarks inside the proton starts to be relevant. At LHC, the associated charm production with W bosons is mainly produced via the diagram shown if Fig. 6.1 ($\bar{s}g \rightarrow W^+ + \bar{c}$ and $sg \rightarrow W^- + c$ processes at the hard scattering level), thus directly probing the strange quark content of the proton.

Strangeness composition in the proton is the worst known light quark composition. Many physical studies would benefit from a better understanding of the proton composition at high energy. In this context, the relevance of the study of the W+c channel, which gives direct access to the proton strange content, is remarkable. In particular, at LHC, such a study would shed light on the strange quark composition over a wide range

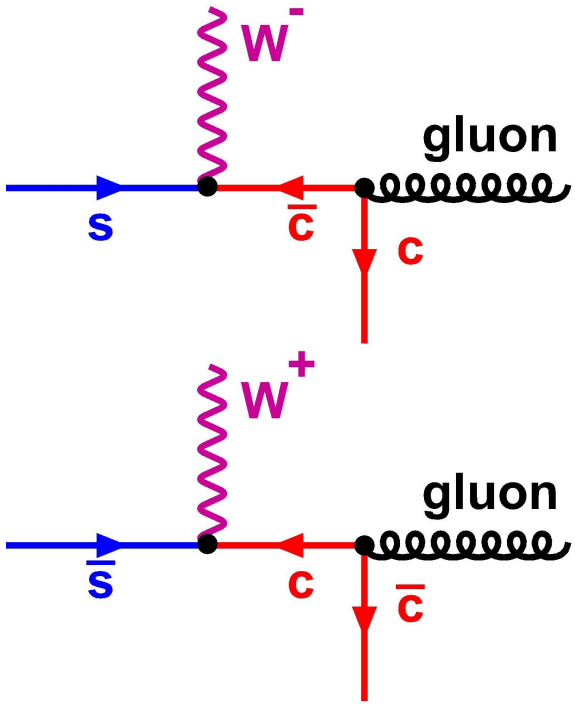

Figure 6.1: Main diagrams at the hard scattering level for associated W - charm production at the LHC.

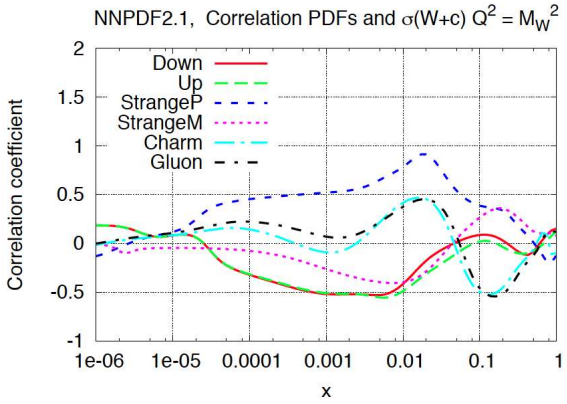

Figure 6.2: Correlation coefficient for the W+c cross section measurement and each of the quarks in the proton (82). The correlation is given for different values of the x parameter.

of the x parameter ($0.0001 \leq x \leq 0.1$), specially at $x \sim 0.02$. In Fig. 6.2 the correlation coefficient of the W+c cross section with each of the parton PDFs is shown. It gives an idea of the sensitivity of this measurement to constrain the different quark proton densities. A high correlation coefficient implies a high sensitivity of this measurement to that PDF. In particular, a big sensitivity to the strange content previously claimed is observed (82).

One of the studies that would specially benefit from such a measurement is the challenging determination of the W boson mass (83). The W boson mass is a key parameter in the Standard Model. This model does not predict the W boson mass, but it predicts the relation between the W boson mass and other experimental observables:

$$M_W = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}} \frac{1}{\sin(\theta_W)\sqrt{1-\Delta r}}$$

where α is the constant coupling, G_F Fermi's constant, θ_W the weak angle and Δr the radiative corrections. These radiative corrections include loop diagrams in which the Higgs boson and the top quark appear. As a consequence, Δr depend on the top quark mass and the Higgs boson mass. Being the current precision on the top quark mass of 0.9 GeV (measured by Tevatron) a precision of 5 MeV on the W boson mass is required in order to have both measurements contributing equally to the Higgs boson mass uncertainty. The world average measurement has an uncertainty of 23 MeV, 5

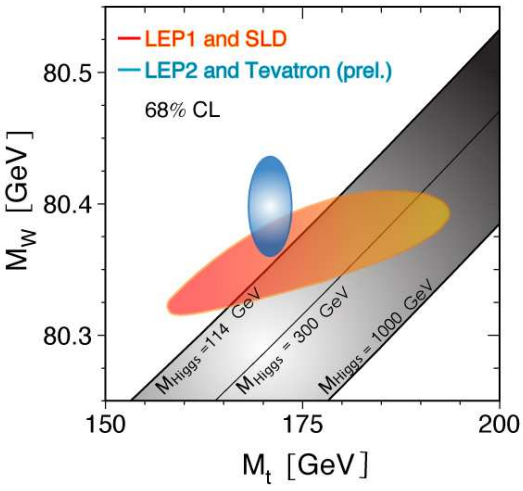

Figure 6.3: World average measurement for the W boson mass and the top quark mass. The dark region represents the allowed region for the SM depending on the Higgs mass.

times bigger than the desired one [1]. In Fig. 6.3 the constrain set by the combined results on the top quark mass and the W boson mass is shown. In this plot the value of the W boson mass versus the top mass is shown. The uncertainty in the top mass and W boson mass measurement is shown as an ellipse centered in the world average value. This average value in this Fig. is shown to be one sigma away from the limit on the Higgs mass imposed by LEP measurements.

The leading theoretical uncertainties on the W boson mass measurements are proton PDFs. A better understanding of the proton PDFs is mandatory to reduce the uncertainty of the W boson mass.

In Fig. 6.4 the differences beyond their current uncertainties on the strange PDF from the different PDF sets, in the range of x 0.01-0.1 can be clearly noticed.

In this Chapter we present the measurement of the ratios $R_c^{\pm} = \sigma(W^+ + charm)/\sigma(W^- + charm)$ and $R_c = \sigma(W + charm)/\sigma(W + \geq 1\ jet)$. The measurement of ratios instead of cross sections has the advantage of being subjected to lower uncertainties (many uncertainties are cancelled when performing ratios, like the luminosity uncertainty). It is also easier since we do not have to deal with certain efficiencies that are also canceled

[1] We don not consider recent results on M_W and Higgs searches

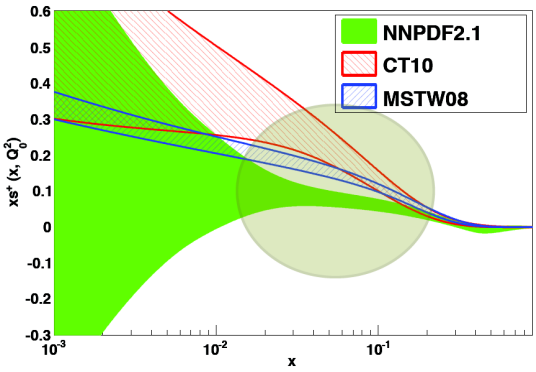

Figure 6.4: Strange quark content (PDF) on the proton as a function of the x parameter for the PDF sets used in this Chapter: MSTW, CT10 and NNPDF. A lack of agreement is observed between the three groups in the intermediate region (0.01-0.1).

out when performing the ratio. At the same time, the strange proton composition is also sensitive to the ratios, supporting this choice. At the LHC, analytical calculations from the MCFM program (44) at next-to-leading order (NLO) predict ratios $R_c^{\pm} \equiv \sigma(\mathrm{W}^+ + charm)/\sigma(\mathrm{W}^- + charm) \approx 1$ and $R_c \equiv \sigma(\mathrm{W}+\mathrm{charm})/\sigma(\mathrm{W}+\mathrm{jets}) \approx 0.1$.

Other contributions to this final state are strongly suppressed. For example the production of a W boson with a charm jet is also possible with a d quark in the initial state ($\mathrm{dg} \rightarrow \mathrm{W}^- + \mathrm{c}$). However this process is Cabibbo suppressed with the matrix element $\|V_{dc}\|$ being small (≈ 0.2). This contribution is even lower for the W^+ ($\bar{\mathrm{d}}\mathrm{g} \rightarrow \mathrm{W}^+ + \bar{\mathrm{c}}$), since the $\bar{\mathrm{d}}$ is not a valence quark. Both contributions account approximately for 15% and 5% of the signal respectively. Gluon splitting of the type $\mathrm{d}\bar{\mathrm{u}} \rightarrow \mathrm{W}^- + \mathrm{g} \rightarrow \mathrm{W}^-\mathrm{b}\bar{\mathrm{b}}$ or $\mathrm{d}\bar{\mathrm{u}} \rightarrow \mathrm{W}^- + \mathrm{g} \rightarrow \mathrm{W}^-\mathrm{c}\bar{\mathrm{c}}$ will also contribute to our background and our signal (respectively), but with a smaller impact ($\approx 1\%$).

Charm quarks are the third most massive quarks of the Standard Model. They have spin 1/2 and charge 2/3. Charmed-hadrons decay (due to conservation of the internal charm quantum number by the strong interaction) via the weak force. In charmed hadrons, the c quark is preferentially transformed into a s quark. Thus, charmed-hadrons decay is mostly observed into kaons and pions. This decay process has a lifetime of the order of 10^{-13} seconds. As a consequence, these hadrons usually travel several mm before decaying and thus, producing a secondary vertex that can be observed in the detector. The presence of the charmed-hadrons can be distinguished statistically

from a b-meson or a light meson for having a short flight distance compared to the b hadrons and larger than the light hadrons. Other properties of the charmed-meson jets that allow us to distinguish them from other kind of jets are their multiplicity, mass and the opening of the jet cone. They are known to be intermediate between the b-hadrons and the light quark hadrons in terms of mass of the jet, number of tracks in the jet and angle of the jet cone. The techniques that allow us to distinguish a charmed-hadrons from other hadrons are known as c-tagging techniques, and exploits the characteristics shown before.

For this analysis we will consider as signal the presence of a charm-tagged jet in addition to a W boson. We will call "W+b" in case the jet is coming from a b quark, and "W+udsg" if we have a light quark, either from u quark, d quark, s quark or a gluon.

6.1 Samples

The data sample used for this measurement is the sample employed for the W cross section determination. It was already presented in the previous Chapter.

Large MC samples of the signal and main background processes are generated. They are used for validation of the employed methods and various checks. They are reported in Tab. 6.1.

W signal ($W\rightarrow \mu\nu$) as well as other EWK processes (such as $Z\rightarrow \mu\mu$, $W\rightarrow \tau\nu$ and $Z\rightarrow \tau\tau$ production) are generated with POWHEG (42). No cuts were applied in phase space at the generator level, except for the Z/γ^* case ($M_{\mu\mu} > 20$ GeV). The PDF set used in these POWHEG productions was CT10.

The analysis requires at least one jet in the final state and sets a rather loose cut on the maximum number of hard jets. Therefore a POWHEG approach, which should correctly describe the inclusive production of W plus one hard jet (plus collinear/soft QCD) is suitable for this analysis. The $c\bar{c}$ and $b\bar{b}$ contributions are small and originate mostly from gluon splitting. There are also higher order processes like $bq \rightarrow bWq'$ than compete with gluon splitting, but both contributions are included in PYTHIA via parton evolution. An important advantage of the POWHEG approach (a NLO generator) is the possibility to perform detailed comparisons with NLO PDFs, which give a more accurate prediction.

Generator	Process	σ (pb)	Events
POWHEG	$W^+ \to \mu^+\nu_\mu$	6152	~2M
POWHEG	$W^- \to \mu^-\bar{\nu}_\mu$	4286	~2M
POWHEG	$W^+ \to \tau^+\nu_\tau$	6152	~2M
POWHEG	$W^- \to \tau^-\bar{\nu}$	4286	~2M
POWHEG	$Z \to \mu^+\mu^-$	1666	~ 2M
POWHEG	$Z \to \tau^+\tau^-$	1666	~ 2M
MADGRAPH	$t\bar{t}$	162	~1.3M
MADGRAPH	Single-top (t-channel)	21.5	~0.5M
PYTHIA	Inclusive μ QCD	79688	28M
PYTHIA	WW	42.9	~2M
PYTHIA	WZ	18.3	~2M
PYTHIA	ZZ	5.9	~2M
MADGRAPH	$W + jets$	6152	~15M

Table 6.1: Summary of analyzed Monte Carlo samples for the various signal and background processes. Reported cross sections correspond to NNLO FEWZ MSTW08NNLO for W and Drell-Yan cases and to NLO CTEQ66 cross sections for the remaining backgrounds. No cuts have been applied at the generator level whenever is possible. Exceptions (QCD, processes with photon exchange) are discussed in the text.

We also use an alternative MadGraph LO approach in the analysis. This allow us to cross-check the POWHEG approach and evaluate explicitly any potential dependence on the details of hard jet production or on the description of heavy-quark contributions associated with higher jet multiplicities in the final state.

Important backgrounds in this analysis are $t\bar{t}$ and single top (the latter being fully dominated by the t-channel $bq \to tq'$ component). Event samples of both are generated with MadGraph (84) interfaced with PYTHIA. QCD ($\hat{p}_T > 20$ GeV, $p_T^\mu > 15$ GeV and diboson backgrounds (WW, WZ, ZZ) are generated with PYTHIA. All PYTHIA generations use the Z2 tune (85) to model the underlying event characteristics. All PYTHIA LO generations use CTEQ6L PDFs.

For the signal ($W \to \mu\nu$) and $t\bar{t}$ backgrounds, we have considered simulations that take into account the presence of pileup events ($\sim 2-3$ minimum bias interactions on top of the hard process) to match the experimental conditions.

Details on the Monte Carlo samples used in this analysis are given in Tab. 6.1.

6.2 Selection

The selection of W+c events is aimed to select a pure sample of W candidates, containing at least a high p_T jet in the event within detector acceptance, removing as much as possible the background arising from other processes delivering events with a similar topology in the final state.

The basic selection consist in finding a W candidate, using the selection described in the previous Chapter. This set of requirements was shown to be optimal in terms of muon reconstruction and signal extraction.

We require some minimal quality criteria to preselect muon candidates using tracks traversing the muon and inner tracking volumes: the muon must be reconstructed by two different reconstruction algorithms: the muon track must contain one or more pixel hits; the number of hits used in the inner tracker (pixel+strips) must be greater than 10 to ensure an accurate p_T computation; the transverse impact parameter of the muon with respect to the beam should be smaller than 2 mm, to reject cosmic background; the χ^2 per degree of freedom of the global muon fit (system-inner tracker) must be less than 10; at least one hit in the muon spectrometer must be included in the final refit and at least two muon stations must be used in the global muon fit.

In addition, some specific criteria to select clean $W \to \mu\nu$ candidates are applied:

- the transverse momentum of the muon track must be larger than 25 GeV,
- and the muon lie in the region $\mid \eta_\mu \mid < 2.1$.
- less than two muons with $p_T^\mu > 10$ GeV, in order to reduce the Drell-Yan contamination,
- the relative combined isolation variable, built from the sum of ECAL, HCAL and inner tracker transverse energies or momenta in a $\Delta R < 0.3$ cone,

  $$I^{\rm rel}_{\rm comb} = \sum (E_T({\rm ECAL}) + E_T({\rm HCAL}) + p_T({\rm tracks}))/p_T(\mu),$$

  must satisfy $I^{\rm rel}_{\rm comb} < 0.1$,
- the missing transverse mass, M_T, built from p_T^μ and the missing transverse energy, $E_{\rm T}^{\rm miss}$, must be larger than 50 GeV,

  $$M_T \equiv \sqrt{2\ p_T^\mu\ E_{\rm T}^{\rm miss}\ (1 - \cos(\phi_\mu - \phi_{E_{\rm T}^{\rm miss}}))} > 50\ {\rm GeV}.$$

Compared with the optimized in the previous Chapter selection we have added the $M_T > 50$ GeV requirement, which strongly suppresses the contribution of the QCD background. Given the nature of the present measurement - we are measuring ratios - and the reasonable agreement between data and simulations for the QCD shape background, this cut allows almost full suppression of the QCD background with a negligible impact on the systematic uncertainty of the measurement.

Jets in the studies presented here are reconstructed using the anti-k_T (86) clustering algorithm with the size parameter R = 0.5. Jets in CMS are reconstructed using the particle flow technique (87, 88).

In addition to the $W \rightarrow \mu\nu$ selection just described, a W + jets sample is selected in the following way:

- at least one jet with $p_T^{jet} > 20$ GeV and $|\eta^{jet}| < 2.1$, separated from the muon from the W decay by $\Delta R > 0.3$,

- less than three jets with $p_T^{jet} > 40$ GeV and $|\eta^{jet}| < 2.1$, in order to reduce top backgrounds.

After this selection, the most important processes that has a similar signature to that of the signal are:

- $t\bar{t}$ events. t quarks decay almost exclusively in a W boson plus a b quark. In the case of a leptonic decay of the W, these events are characterized by two high p_T leptons and two b jets. In the hadronic decay, it could contain up to 6 jets of different flavour. At least two of these jets contain a b-quark (long decay length, high invariant mass, not as collimated as a light jet). $t\bar{t}$ background is partially removed by restricting to events with a reduced number of energetic jets.

- Single-top events. This channel is dominated by the t-channel production ($bq \rightarrow tq'$). Thus, it is characterized by a t quark accompanied by another quark in the final state. The t quark decays into a b and a W (as shown before). As a consequence, this channel is characterized by a W and two quark jets, one of them containing a b quark. Setting strong constraints on the number of jets or in the mass of the most energetic jet reduce the contribution coming from this process.

- W+b($\bar{\text{b}}$) events. This irreducible background has a similar signature to that of the signal. The properties that we can exploit to distinguish it from the signal were shown before and uses the properties of the massive b quark jet.

- W+light-jet events. This is the dominant background. Heavy flavour tagging techniques, exploiting specially the flight distance, allow us to separate this process from the signal.

- QCD multijets events. This huge background is largely removed by the W selection described in Chapter 6.Setting an additional requirement on the transverse mass, M_T, of the MET-muon system to be greater than 50 GeV reduces this background to a negligible level.

- other electroweak processes (such as Drell-Yan or dibosons). The basic selection (W candidate with a high p_T jet) already reduces this contribution to reasonable levels, not requiring dedicated selection requirements.

The two first backgrounds ($\text{t}\bar{\text{t}}$ and single-top) are reduced by a loose cut on the number of high energy jets. The remaining top background will be separated from the signal using the long flight distance that b quark jets have. In Fig. 6.5 the number of jets after the W selection for data, W+c, single-top and $\text{t}\bar{\text{t}}$ from MC simulation is shown. The high jet multiplicity of $\text{t}\bar{\text{t}}$ events can be noticed in this plot. To remove this background we veto events with three or more selected jets with energy over 40 GeV.

No additional criteria are applied to reduce the W+b background due to its similarity to the signal. We will rely on the MC prediction and assume a 100% uncertainty in the W+b normalization.

W+light jets background is the main contributor to the sample. In order to distinguish it from the W+c signal, b-tagging techniques are used, as it will be explained in the section devoted to the method developed.

With this selection criteria, we select 31730 events. As evaluated by MC, the non W+jets background amounts to 5.3%, and it is composed by Drell-Yan (2.1%), $W \to \tau\nu$ (1.4%) and top events (1.4%). Subtracting these backgrounds according to the Monte Carlo expectations we obtain a corrected W+jets yield of:

$$N(W + jets; p_T^{jet} > 20 \text{ GeV}, |\eta^{jet}| < 2.1) \quad = \quad 30038.0 \pm 178.1 \ (stat.)$$

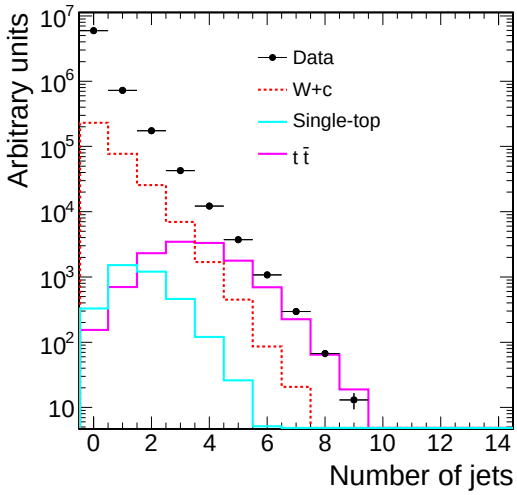

Figure 6.5: Number of jets in the event (within pseudorapidy and p_T acceptance defined in this analysis) after the W selection.

where the uncertainty is purely statistical.

Fig. 6.6 shows the M_T distribution for W + $\geq$ 1 jet selected events, separately for W^+ and W^-. Good agreement between total MC prediction and data is observed. Fig. 6.7 shows the transverse momentum distribution of the most energetic jet in the event compared with POWHEG and MadGraph references. They are in agreement with expectations (MadGraph predictions are more accurate in the high-p_T tail). Similar conclusions can be extracted from the jet multiplicity distributions in Figs. 6.8. POWHEG correctly predicts the kinematics in the events with 1 or 2 jets, failing at high multiplicities.

6.2.1 Definition of W+c, W+b and W+ligh quark components at the generator level

The definition of the jet flavour is an important part in the analysis to correctly define the measured ratio. W plus jets events in CMS usually contain quarks of different flavours at the same time. The tagging of jet as a certain flavour jet will depend on our definition, but the result will be independent of such a definition provided we stay coherent with this definition through the whole analysis.

The definition adopted in this analysis consider an event in Monte Carlo as "W+b"

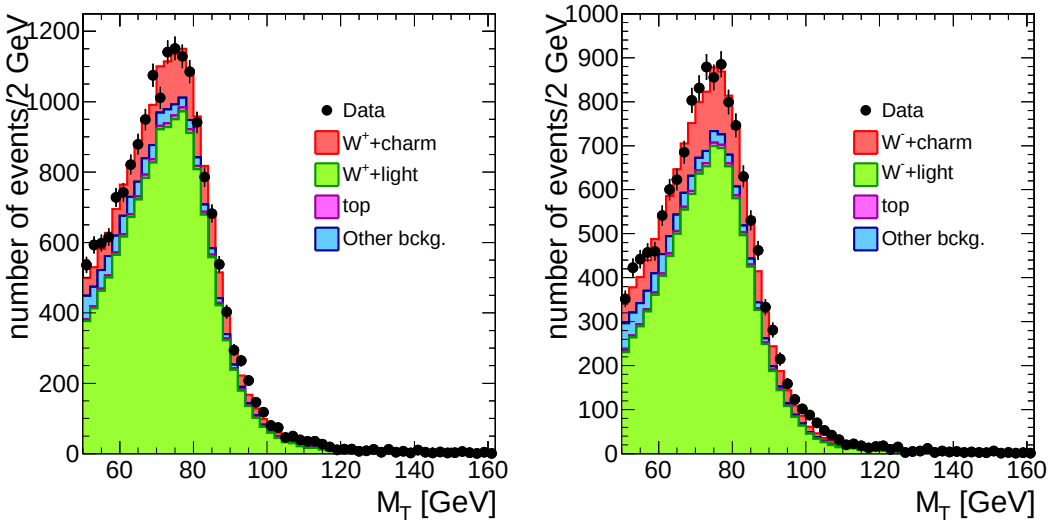

Figure 6.6: Distribution of the transverse mass in selected W^++jets and W^-+jets events compared with POWHEG predictions after the selection. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contributions of the other backgrounds. Histograms are stacked.

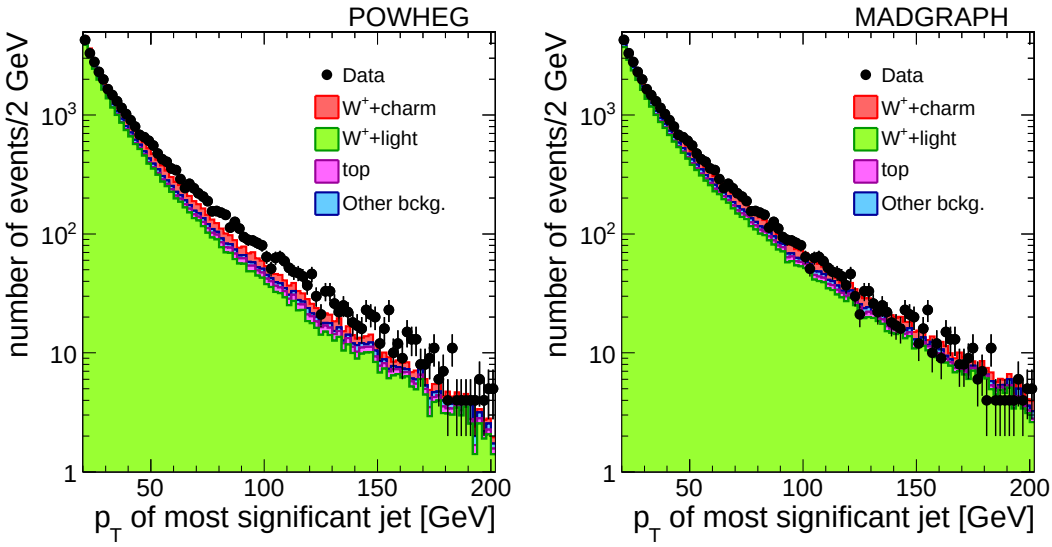

Figure 6.7: Distribution of the transverse energy of the most significant jet ($p_T^{jet} > 20$ GeV) in selected W+jets events compared with POWHEG (left) and MadGraph predictions (right) or the signal. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contribution of the other backgrounds. Histograms are stacked.

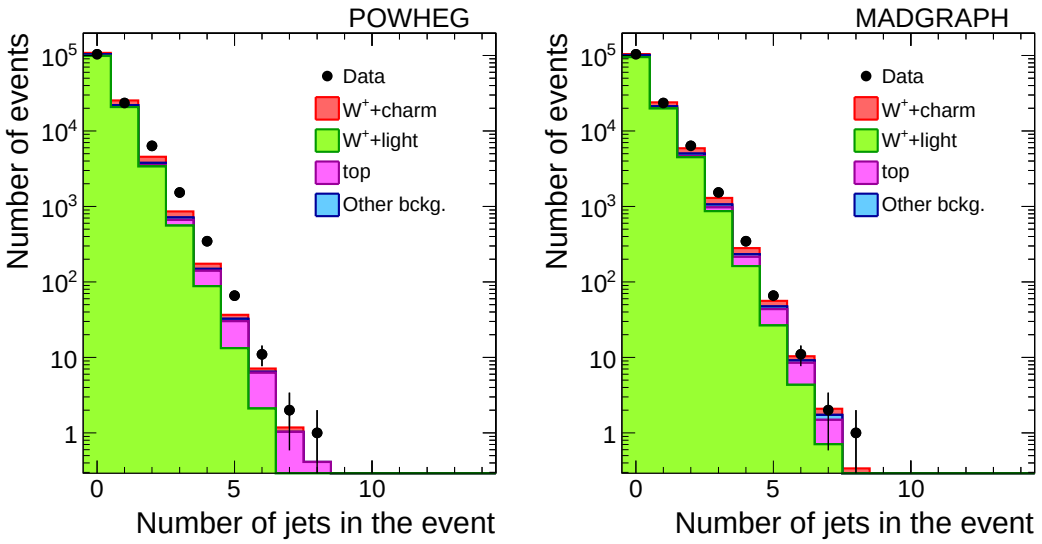

Figure 6.8: Distribution of the number of jets with $p_T^{jet} > 20$ GeV in selected W$^+$+jets events compared with POWHEG (top) and MadGraph predictions (bottom). In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contribution of the other backgrounds. Histograms are stacked.

if it has a b (or $\bar{\text{b}}$) entry at generator level with $p_T > 10$ GeV and $|\eta| < 2.5$, "W+c" if it has a c (or $\bar{\text{c}}$) entry with $p_T > 10$ GeV and $|\eta| < 2.5$, and "W+udsg" otherwise. This is a way to ensure that tagged jets with $p_T > 20$ GeV and $|\eta| < 2.1$ (the cuts that will be used in the analysis described in the next sections) are taken into account properly as heavy quarks despite fluctuations in energy and angle with respect to the generated values. The definition also ensures that there is a reasonable transverse momentum for a charm candidate to be considered for tagging purposes. We avoid for instance calling W+c an event that has a charm entry with almost zero p_T and one additional hard gluon jet. Accepting an event like this after lifetime tagging cuts is unlikely, but considering it as a charm event would be detrimental for the analysis because it would contribute to the tagging region dominated by light quarks. Let us note that this convention is mostly a matter of classification. Experimental cuts ensure that the selected events are constrained kinematically to the region where jets have $p_T > 20$ GeV and $|\eta| < 2.1$. Changing the definition of the W components modifies minimally the results of the fitted charm fractions and should not produce any bias in the final measurements, simply because the efficiencies used in ratios is corrected according to our definition.

This classification is applied in Fig. 6.6, Fig. 6.7 and Fig. 6.8 to differenciate between the contributions. A clear W+c component can be seen in these Figs.

6.2.2 W+charm extraction

Heavy flavour jets (with c or b quarks) are characterized by a long flight distance. In order to enhance the W + charm content in the sample we require the presence of a secondary vertex with two or more associated tracks, coming from the decay of the charmed or b-hadron, significantly displaced from the primary interaction vertex, following the "simple secondary vertex high-efficiency" (SSVHE) algorithm of CMS (89) (in the HE, High Efficiency, version we require at least 2 tracks in the vertex, as opposed to the HP, High Purity, version with a minimum of 3 tracks in the vertex). Only vertices with decay length uncertainties smaller than 1.5 mm are used.

A lifetime tagging requirement significantly enhances the fraction of W+charm events in the final state. The presence of W+b and top backgrounds is also enhanced but, due to absence of a significant b component in W+jets subsample and the moderate top backgrounds, the W+c signal is clearly dominant after this last requirement. Instead of using directly the decay length, the following discriminant D is defined as $D = \text{sign(S)}\ \log(1 + \text{abs(S)})$, where S is the decay length significance. The logarithmic function is used in order to have more sensitivity with lower values of the discriminator (light contribution). Jets without a secondary vertex or not significant enough (an intrinsic requirement of ≈ 1 is set on the D value of the jet secondary vertex) do not enter in the plot. Negative vertices are also included in the analysis. The (small) number of events with negative vertices is due to the tails of tracker resolution (and possibly jet angular resolution). The definition of a positive and negative vertex are shown in Fig. 6.10 and Fig. 6.11 respectively. Negative vertices are built using the same criteria and cuts as the positive ones, but for cases in which the vector connecting primary and secondary vertex is opposite to the jet direction. Fig. 6.9 shows the distribution of the lifetime discriminant variable for the most significant (the one with largest value of the discriminant) jet in $W^+ \to \mu^+\nu_\mu \geq +1$ jet and $W^- \to \mu^-\bar{\nu}_\mu \geq +1$ jet selected samples. A clear W+charm component is observed, in good agreement with Monte Carlo predictions.

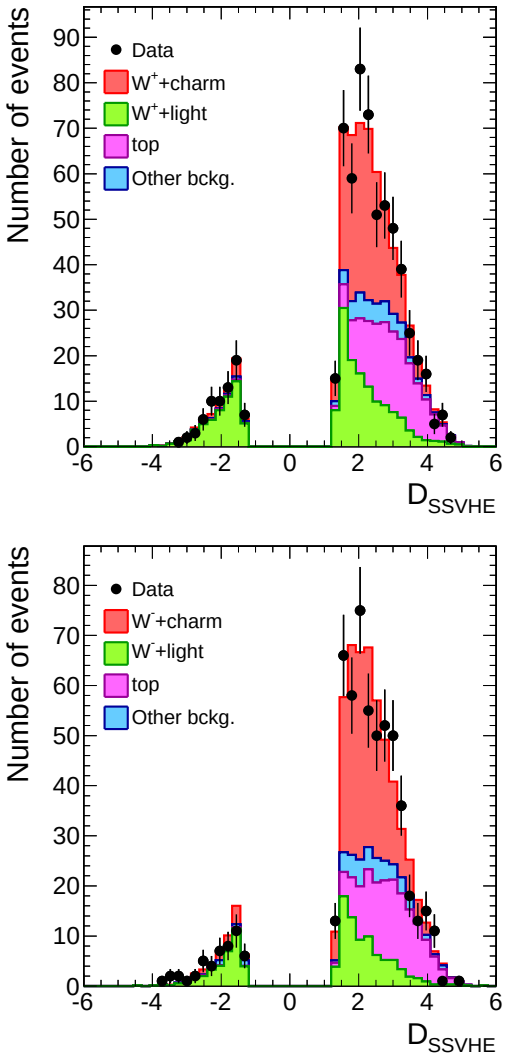

Figure 6.9: SSVHE discriminator in selected $W^+ \rightarrow \mu^+\nu_\mu + \geq 1$ jet (top) and $W^- \rightarrow \mu^-\bar{\nu}_\mu + \geq 1$ jet (bottom). In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contribution of the other backgrounds. Histograms are stacked.

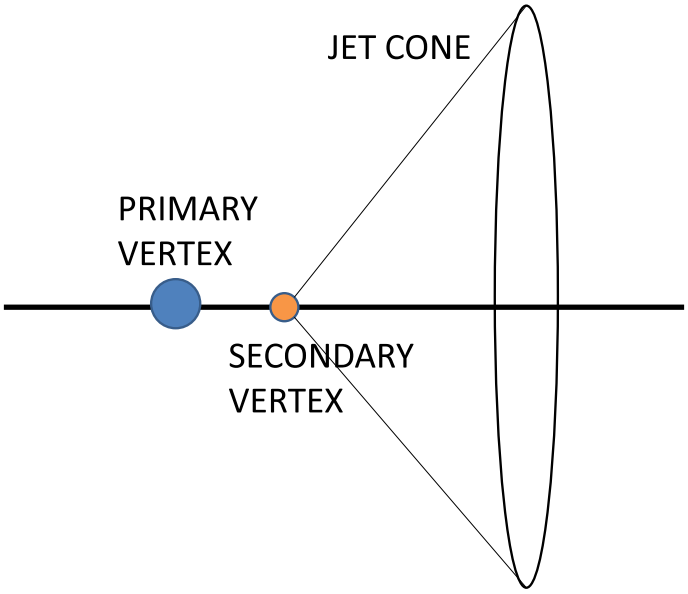

Figure 6.10: An example of a positive secondary vertex.

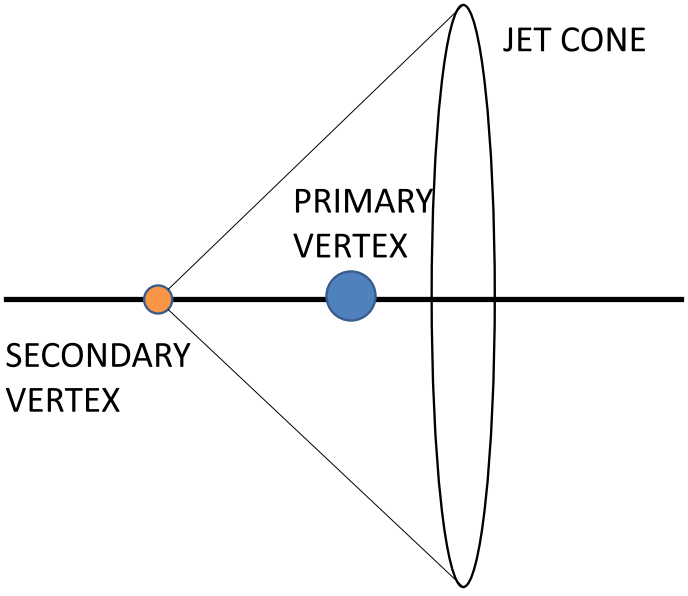

Figure 6.11: An example of a negative secondary vertex.

Sample	Muon	$M_T > 50$ GeV & DY veto	≥ 1 jet $p_T^{jet} > 20$ GeV, $\lvert\eta^{jet}\rvert < 2.1$	< 3 jets with $p_T^{jet} > 40$ GeV	b-tag SSVHE	$\Delta_{\mathrm{decay-length}} < 0.15$ cm
Events with $\mu^{\pm}$	195743	136008	32185	31735	1421	1199
Events with μ^{+}	112292	80259	18237	18010	768	636
Events with μ^{-}	83451	55749	13948	13725	653	563

Table 6.2: Data reduction at every step of the selection process. The number of events is given for the whole muon data sample, and also separated by the muon charge. The Muon requirement means a good quality and isolated ($I^{\mathrm{rel}}_{\mathrm{comb}} < 0.10$) muon with $p_T > 25$ GeV and $|\eta| < 2.1$, according to the criteria described in the text.

Tabs. 6.2, 6.3 and 6.4 show the cumulative evolution of the number of selected events as a function of the applied cuts in data and MC. The final sample composition is detailed in 6.5.

6.3 Measurement of the $\sigma(\mathrm{W}^+ + charm)/\sigma(\mathrm{W}^- + charm)$ ratio

The W+charm signal is extracted from a maximum likelihood fit to the different components of the high-efficiency simple secondary vertex b-tagging variable (SSVHE) distribution.

As reported in the previous section, after the selection process the dominant backgrounds with positive discriminator to W+c are single-top and $\mathrm{t\bar{t}}$ with a small contribution from W plus light-quark/gluon jets. Consequently we perform a fit to determine the normalization of the three components: 1) W+c signal, 2) top backgrounds, 3) light-jets. Shapes are taken from MC. The contribution from the remaining components, including the W+b, is taken from simulation and added to a fourth category "Others". Fig. 6.12 shows the D_{SSVHE} distribution of the different components normalized to the MC expectations. Their shapes are rather different, thus justifying the approach of fitting these three main contributions.

In order to be rather insensitive to the light-jet component, the fit takes also into account negative SSVHE discriminators. Including them in the fit is an effective way to constrain the number of positive vertices from light-quark/gluon sources. We note that the amount of negative light tags is slightly smaller than the amount of positive

Sample	Muon	$M_T > 50$ GeV & DY veto	≥ 1 jet $p_T^{jet} > 20$ GeV, $\lvert\eta^{jet}\rvert < 2.1$	< 3 jets with $p_T^{jet} > 40$ GeV	b-tag SSVHE	$\Delta_{\mathrm{decay-length}} < 0.15$ cm
$W^+ + c$	4560 ± 22	3931 ± 20	2166 ± 15	2159 ± 15	237 ± 5	210 ± 5
$W^+ + udsg$	81406 ± 73	73371 ± 71	14458 ± 38	14442 ± 38	187 ± 4	127 ± 4
$W^+ + b$	295 ± 6	249 ± 5	137 ± 4	136 ± 4	23 ± 2	21 ± 1
$t\bar{t}$	382 ± 1	203 ± 1	202 ± 1	99 ± 1	69 ± 1	66 ± 1
single-t	90 ± 1	61 ± 1	57 ± 1	53 ± 0	32 ± 0	31 ± 1
$W^+ \to \tau^+\nu_\tau$	2551 ± 16	1563 ± 13	249 ± 5	249 ± 5	7 ± 1	6 ± 1
$Z \to \mu^+\mu^-$	16546 ± 18	2205 ± 8	351 ± 3	350 ± 3	12 ± 1	9 ± 1
$Z \to \tau^+\tau^-$	492 ± 4	75 ± 1	24 ± 1	23 ± 1	1 ± 1	1 ± 1
QCD	4166 ± 20	72 ± 3	54 ± 2	52 ± 2	6 ± 1	5 ± 1
Total	110487 ± 82	81730 ± 76	17698 ± 41	17563 ± 41	573 ± 7	475 ± 6

Table 6.3: Expected number of events after every step of the selection criteria, for the several physics processes producing a signal as defined in Section 4. Projection for an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 36$ pb^{-1}.

Sample	Muon	$M_T > 50$ GeV & DY veto	≥ 1 jet $p_T^{jet} > 20$ GeV, $\lvert\eta^{jet}\rvert < 2.1$	< 3 jets with $p_T^{jet} > 40$ GeV	b-tag SSVHE	$\Delta_{\mathrm{decay-length}} < 0.15$ cm
$W^- + c$	4583 ± 18	3949 ± 17	2191 ± 13	2187 ± 13	243 ± 4	212 ± 4
$W^- + udsg$	53720 ± 50	49337 ± 49	10197 ± 26	10186 ± 26	136 ± 3	93 ± 3
$W^- + b$	207 ± 4	178 ± 4	96 ± 3	95 ± 3	18 ± 1	17 ± 1
$t\bar{t}$	383 ± 1	205 ± 1	205 ± 1	100 ± 1	70 ± 1	67 ± 1
single-t	142 ± 1	97 ± 1	92 ± 1	84 ± 1	52 ± 1	49 ± 1
single-t	47 ± 1	32 ± 1	31 ± 1	28 ± 1	17 ± 1	16 ± 1
$W^- \to \tau^-\overline{\nu}$	1880 ± 12	1140 ± 9	188 ± 4	187 ± 4	8 ± 1	6 ± 1
$Z \to \mu^+\mu^-$	16175 ± 18	1924 ± 7	310 ± 3	309 ± 3	11 ± 1	8 ± 1
$Z \to \tau^+\tau^-$	467 ± 4	70 ± 1	23 ± 1	23 ± 1	1 ± 1	1 ± 1
QCD	3989 ± 20	65 ± 3	48 ± 2	47 ± 2	5 ± 1	4 ± 1
Total	81451 ± 61	56901 ± 53	13288 ± 30	13163 ± 30	509 ± 6	424 ± 5

Table 6.4: Expected number of events after every step of the selection criteria, for the several physics processes producing a signal as defined in Section 4. Projection for an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 36$ pb^{-1}.

Source	$N_{bg}/(N_{sg}+N_{bg})$	N_{bg} in 36 pb^{-1}
W+ udsg	$24.5 \pm 0.6\%$	220 ± 4
W+ b	$4.1 \pm 0.2\%$	37 ± 2
$t\bar{t}$	$14.8 \pm 0.2\%$	133 ± 1
single-t	$5.2 \pm 0.1\%$	47 ± 1
$W \to \tau\nu$	$1.3 \pm 0.1\%$	12 ± 1
$Z \to \mu^+\mu^-$	$1.9 \pm 0.1\%$	17 ± 1
$Z \to \tau^+\tau^-$	$0.1 \pm 0.1\%$	1 ± 1
QCD	$1.1 \pm 0.1\%$	10 ± 1
Total Bckg.	$53.0 \pm 0.7\%$	476 ± 5
W + c signal	$47.0 \pm 0.8\%$	423 ± 6

Table 6.5: Final sample composition from MC predictions. The second column shows the fraction of each of the background processes that contributes to the candidate sample. The third column gives the number of events expected of each type in a 36 pb^{-1} sample, based on the theoretical predicted cross sections. The last row gives the prediction for the signal process.

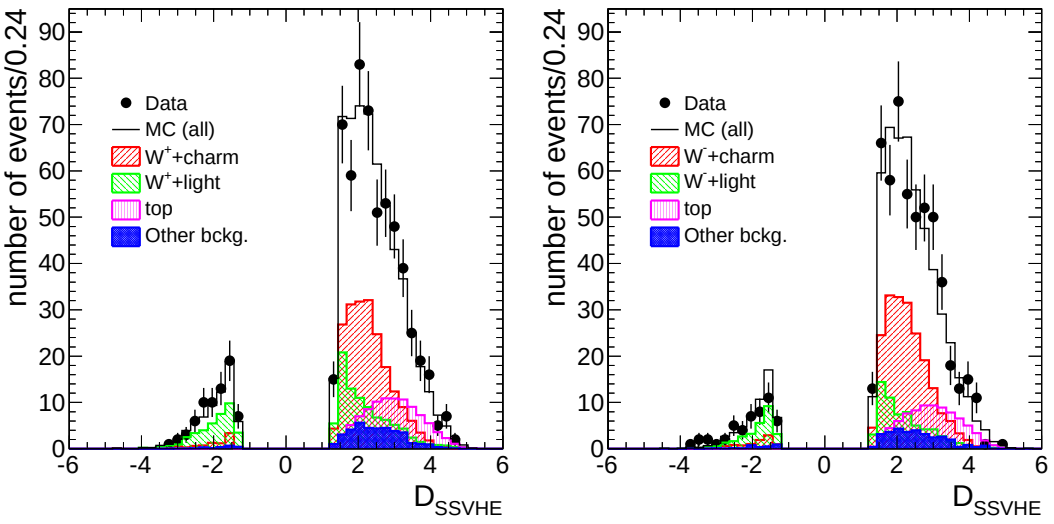

Figure 6.12: SSVHE discriminant distributions of the main components of the enriched W+c sample (separately for W^+, left, and W^-, right), normalized to the expected luminosity and taking into account the cross sections predicted by Monte Carlo. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contributions of the other backgrounds.

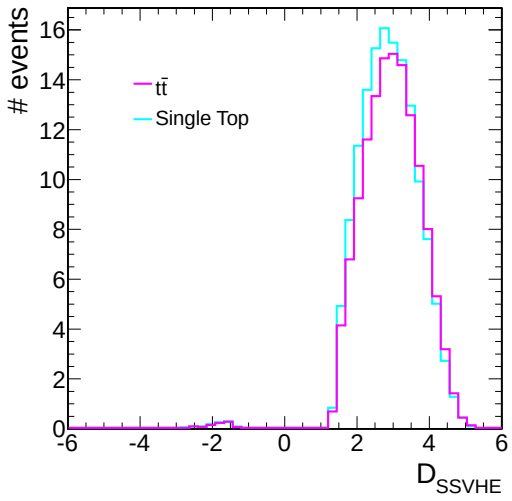

Figure 6.13: Comparison between single-top and $t\bar{t}$ templates in Monte Carlo. The normalization in this plot is arbitrary.

light tags, due to remaining long-lived K^0_S and Λ decays.

The shape of the discriminant variable for single-top and $t\bar{t}$ background events is similar (Fig. 6.13), and it is taken from Monte Carlo. As a consequence, we can add both contributions and take them in the fit as a single component (top background contribution). Due to the dominance of single-top in t-channel, top backgrounds will be larger in $W^+ +$ charm than in $W^- +$ charm (in absolute value).

The results of the fits to the $W^+ + charm$ and $W^- + charm$ distributions are shown in Fig. 6.14. We obtain the following yields:

$$\begin{aligned} N(W^+ + charm) &= 247.4 \pm 40.2\ (stat.) \\ N(W^- + charm) &= 268.9 \pm 38.5\ (stat.) \end{aligned}$$

Dividing both numbers, we get the following measurement of the W + charm charge ratio:

$$R_c^{\pm}(p_T^{jet} > 20\ \text{GeV}, |\eta^{\text{jet}}| < 2.1) = 0.92 \pm 0.20\ (stat.)$$

Possible sources of systematic uncertainties have small contributions to this ratio. They are discussed in Section 6.5.

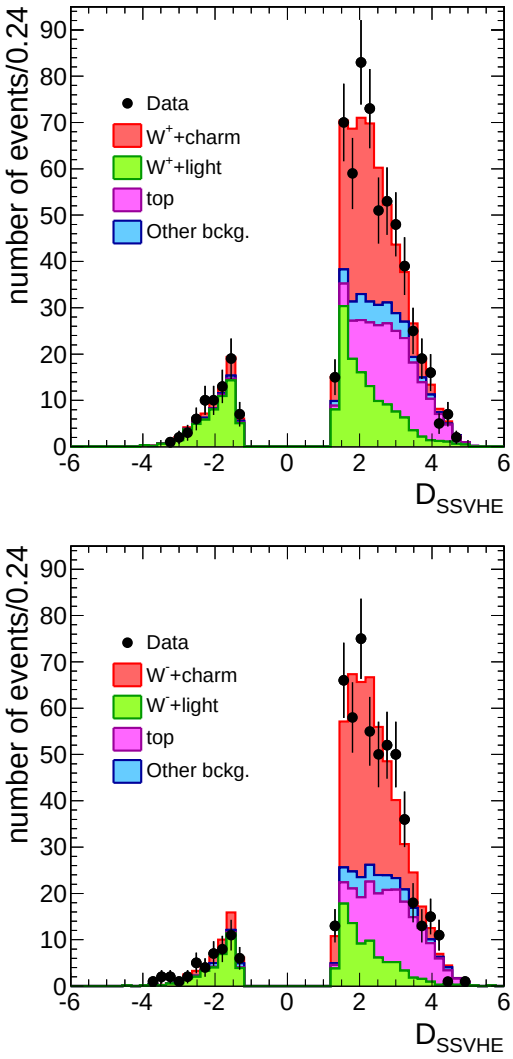

Figure 6.14: Top: fit to the SSVHE discriminator of the $\mathrm{W}^+ + charm$ selected sample. Bottom: fit to the SSVHE discriminator of the $\mathrm{W}^- + charm$ selected sample. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the other contributions of the backgrounds. Histograms are stacked.

6.4 Measurement of the $\sigma(\mathrm{W}+\mathrm{charm})/\sigma(\mathrm{W}+\mathrm{jets})$ ratio

The yields obtained in the previous section and the overall W+jets yields obtained before are combined to obtain a measurement of the W+charm over W+jets ratio for $p_T^{jet} > 20$ GeV. We use the expression:

$$R_c(p_T^{jet} > 20 \text{ GeV}, |\eta^{\text{jet}}| < 2.1) \quad = \quad \frac{N(W^+ + charm) + N(W^- + charm)}{\epsilon_c \; N(W + jets)}$$

where ϵ_c is a correction factor determined in Monte Carlo as:

$$\epsilon_c \equiv N_{selected\ and\ tagged\ jets}/N_{total\ selected\ jets},$$

according to the phase space definition for the theoretical prediction introduced in Section 6.2.1. This correction factor coincides with the SSVHE tagging efficiency in the limit in which all other selection criteria have the same effect on W+c and W+jets and both samples have similar kinematic properties. Effectively, ϵ_c also absorbs remaining kinematic effects and also acceptance corrections due to the mismatch between the reconstructed jet parameters and the properties of the parton that initiates the jet. Given the reasonably accurate description of the tracker response in CMS simulations the central value for this efficiency is taken from Monte Carlo with an uncertainty determined from dedicated studies (section 6.5).

Using the Monte Carlo estimate $\epsilon_c = 0.1218 \pm 0.0019$ $(stat.)$ we obtain:

$$R_c(p_T^{jet} > 20 \text{ GeV}, |\eta^{\text{jet}}| < 2.1) \quad = \quad 0.141 \pm 0.015 \ (stat.)$$

where the uncertainty is statistical.

6.5 Systematic uncertainties

All the systematic uncertainties are computed by comparing the results shown before with the ratios computed by applying the modifications suggested in each of the systematic errors considered. The following sources of systematic uncertainty are considered in the analysis:

- **Difference in efficiency between μ^+ and μ^-.** This source is only relevant for the $R_c^{\pm}$ analysis. It has been computed in the CMS $\mathrm{W} \to \mu\nu$ analysis and amounts to $\leq 1\%$.

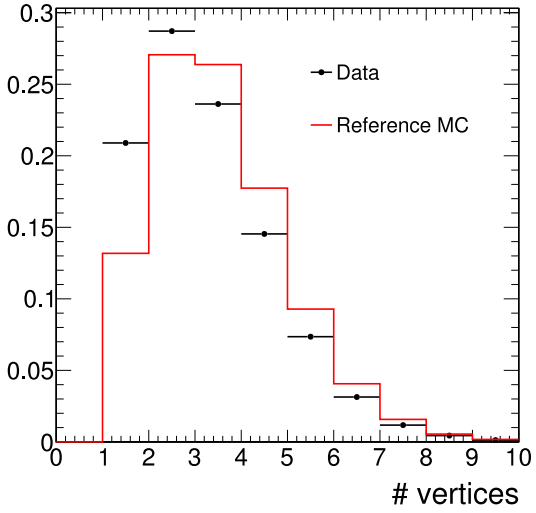

Figure 6.15: Distribution of the number of primary vertices in data and MC simulation used.

- **Muon momentum scale and resolution**. This is computed by distorting the muon resolution in Monte Carlo in order to match the performance found in data following the procedure shown in Chapter 4 (SIDRA).

- **Pileup.** The simulated samples used do not fully reproduce the real pileup conditions of the 2010 run. MC distributions are modified reweighting the events according to the differences between data and pileup simulation for the distribution of the number of primary vertices (Fig. 6.15).

- **Jet resolution and jet scale**. We change the jet scale resolution by 3%, as indicated by dedicated jet resolution studies (90). In the case of heavy quarks the uncertainty is increased to 8% to account for unknown flavor-specific differences. This should also cover possible uncertainties on the jet resolution (a fraction of $\sim 15\%$ (91)) .

- **Jet multiplicity**. The POWHEG Monte Carlo is not expected to reproduce so accurately the final states with high hard jet multiplicity (≥ 2). The analysis only uses the lifetime tag of the most significant jet and signal events with two c or b-jets (via gluon splitting) are also reasonably well described. Therefore a small

uncertainty due to this source is expected. It is anyhow estimated by reweighting W+jet events in the POWHEG Monte Carlo in order to match the jet multiplicity distribution observed in MadGraph (Fig. 6.8). This is done independently for light and heavy quarks.

- **Tracking resolution**. Fig. 6.16 shows the distribution of the number of tracks associated to the secondary vertex in the W+c selected sample. There is reasonable agreement with the MC predictions. In order to check the sensitivity to differences in tracking reconstruction between data and MC, we distort the MC predictions in order to provide the optimal agreement with data. We define a probability P for a track to be missed on top of the MC vertex reconstruction. This probability is defined uniformly, i.e. we assume that the probability to miss a track is independent of the number of charged particles originally produced at the vertex, the decay length and the uncertainty on the decay length. With this scheme there is a $N_V \times P^i \times (1-P)^{(N_V - i)}$ probability for a MC vertex with N_V to become a MC vertex with $(N_V - i)$ tracks. The resulting MC distribution is normalized to the total number of vertex entries in data and a χ^2 of the differences between data and the distorted MC predictions (for the vertex track multiplicity distribution) is defined. We found an optimal matching for $P = 8\%$ and use the corresponding distorted MC to assign a systematic uncertainty due to tracking resolution. Fig. 6.16 also shows a more extreme case ($P = 14\%$) which is clearly inconsistent with the observed distribution in data.

- **Uncertainties in top templates**. This is estimated as the difference in the resulting ratios using a data-driven template and the top Monte Carlo template. The data-driven template is built using a top-enriched control sample. We use the same selection but inverting the requirement on the number of high energy jets (i.e. the control sample is composed of events with three or more jets with $p_T^{jet} >$ 40 GeV). Fig. 6.17 shows a comparison of the data-driven templates obtained in data and in the top Monte Carlo, together with the expected Monte Carlo shape after the full W+c selection. While the MC control sample shows a slightly larger tagging discriminator that the MC in average, due to the presence of more b-jets in average, the data control samples has a softer discriminator spectrum, due to a remaining W+jet background. The maximum difference between the

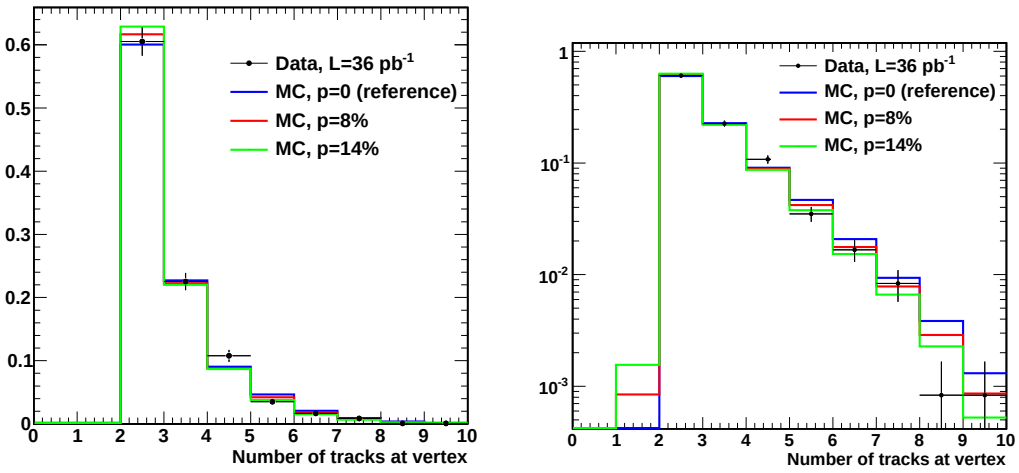

Figure 6.16: Distribution of the number of tracks associated to the secondary vertex for the W + charm selected events both in data (points) and MC (blue solid line), in linear (left) and logarithmic (right) scales. The MC prediction is modified including an additional probability to miss a track in the association to a vertex. The cases $P = 8\%$ and $P = 14\%$ to miss a track are shown. The best agreement with data is reached with $P = 8\%$.

results obtained with Monte Carlo and the results obtained with the data-driven template is conservatively considered as the systematics due to the modeling of top backgrounds.

- **Uncertainty in the light-quark contribution**. This is conservatively estimated by assuming that the excess of positive light-quark tags with respective to negative light-quark tags has a 20% uncertainty. This variation absorbs effectively discrepancies in the number of K^0_S, Λ and conversions observed in dedicated CMS studies (92).

- **W+b contribution**. The W+b contribution is changed by 50%, covering the variations observed between PYTHIA and MadGraph predictions for that component.

- **Other backgrounds taken from MC**. We estimate a 10% variation in the Drell-Yan backgrounds and a 50% variation in the QCD background (1.5 is the typical size of the k-factor that has to be applied to the LO predictions from

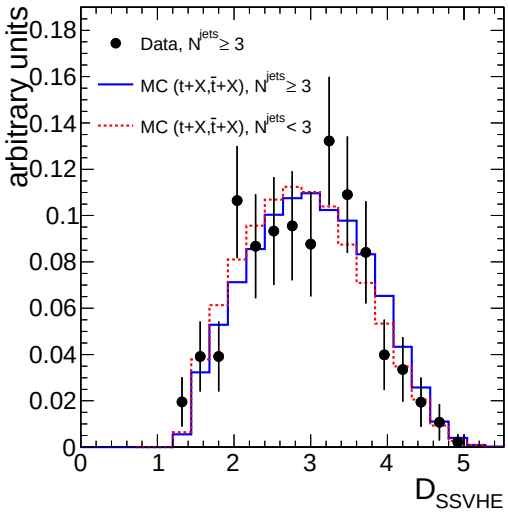

Figure 6.17: Comparison of data-driven templates for top background in data and Monte Carlo. They are also compared with the Monte Carlo truth expectations after the selection process. All templates are normalized to one.

PYTHIA in order to get agreement with the high-p_T inclusive rate observed in multijet events, as seen in the previous Chapter).

- **PDF uncertainties**. Since it is observed that the uncertainties related with strange PDFs within the same PDF set are smaller or equal than the differences between the central values obtained with CT10 (93), MSTW08 (94) and NNPDF21 (95), we assign as PDF systematic uncertainty due to the incomplete knowledge of the PDFs the largest difference obtained using either MSTW08 or CT10 in the case of the charged ratio $R_c^{\pm}$, NNPDF21 or CT10 for R_c.

- **Average energy fraction in charm fragmentation**. We change the average energy fraction carried by the D^* (or D if no excited-D is present) by 2%, which safely covers current estimates for this parameter from previous LEP experiments.

- **Change in charm fragmentation fractions** $BR(c \to D^0, D^{\pm}, D_s, \Lambda_c)$. We change them according to the uncertainties reported in (96). The size of this uncertainty is also consistent with the differences observed between the values present in the CMS Monte Carlo and the central values from reference (96).

Table 6.6: Relative systematic uncertainties (%) in the measurement of $R_c^{\pm}$.

Source	Relative systematics (%)
Charge asymmetry in efficiency	1.0
Muon resolution	<0.1
Pile-up effects	1.8
Jet scale/resolution	1.1
Jet multiplicity	0.7
Tracking resolution	0.3
Top templates in fit	1.4
Light-quark contribution in fit	1.1
W+b background	0.2
Other Monte Carlo backgrounds	1.4
PDF uncertainties	2.2
Charm fragmentation function	<0.1
Charm fragmentation BRs	0.1
TOTAL	4.0

The breakdown of the different systematic contributions to the R_c and $R_c^{\pm}$ measurements is shown in Tabs. 6.6 and 6.7, respectively.

The precision achieved for $R_c^{\pm}$ (a 4% uncertainty) is already at a low level. The larger uncertainty comes from the PDFs. For the R_c value, the total uncertainty amounts to 21%, being driven by the tracking uncertainties and the top background modelling.

6.6 Additional tests

In order to tests the robustness of the measurement, we have repeated the analysis in two different ways. First we have used a different tagging variable (track-counting high-efficiency). Then we also compare the reference method with the results obtained when using MadGraph for the W+jets simulation.

Table 6.7: Relative systematic uncertainties (%) in the measurement of R_c.

Source	Relative systematics (%)
Muon resolution	0.7
Pile-up effects	2.5
Jet scale/resolution	2.3
Jet multiplicity	2.5
Tracking resolution	14.1
Top templates in fit	14.5
Light-quark contribution in fit	3.3
W+b background	2.4
Other Monte Carlo backgrounds	0.2
PDF uncertainties	0.2
Charm fragmentation function	0.2
Charm fragmentation BRs	0.2
TOTAL	21.1

6.6.1 Measurements with an alternative lifetime tagging variable

As a cross check of the analysis, the measurement is also done employing an alternative lifetime discriminator, the so called "track counting high-efficiency" (TCHE) discriminant (89). In this case the discriminant variable is defined as the signed impact parameter significance of the second most significant track in the jet. The $W^+ \to \mu^+\nu_\mu$ and $W^- \to \mu^-\bar{\nu}_\mu$ distributions for this variable are shown in Fig. 6.18.

The study follows exactly the same logic used in previous sections, but in this case the fit is restricted to the region $D_{TCHE} > 3$ and $D_{TCHE} < 20$, in order to suppress the light-quark contribution to almost negligible levels. Actually, the light-quark contribution is fixed in the fit (so it does not contribute to increase the statistical uncertainty of the fit). We obtain the following yields:

$$N(W^+ + charm) = 290.8 \pm 33.0\ (stat.)$$
$$N(W^- + charm) = 298.8 \pm 30.2\ (stat.)$$

Dividing both numbers, we get the following measurement of the W + charm charge

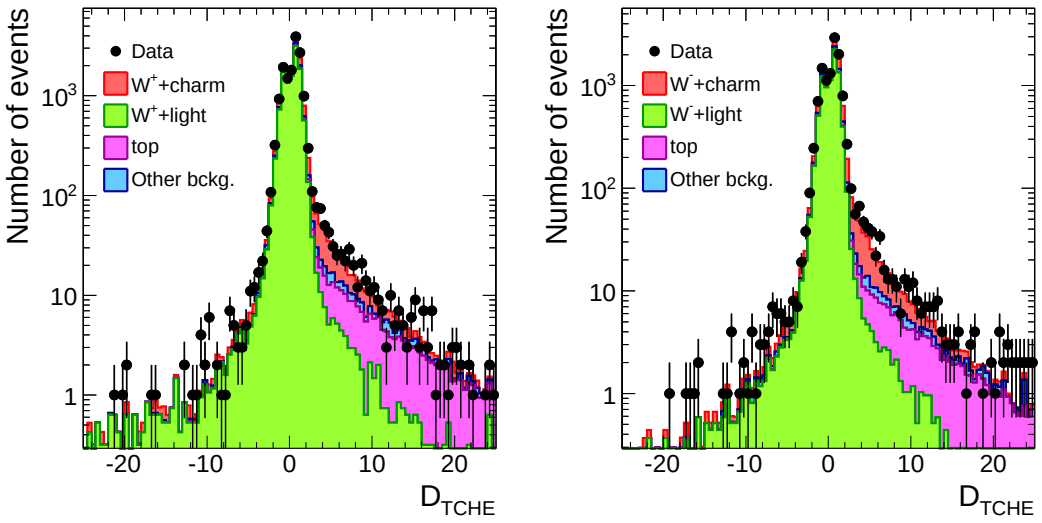

Figure 6.18: TCHE discriminator in selected $W^+ \to \mu^+\nu_\mu + \geq 1jet$ (left) and $W^- \to \mu^-\bar{\nu}_\mu + \geq 1jet$ (right) events. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the other contribution of the other backgrounds. Histograms are stacked. The W+charm signal populates the region $1 < discriminant < 10$.

ratio:

$$R_c^{\pm}(p_T^{jet} > 20 \text{ GeV}, |\eta^{\text{jet}}| < 2.1) \quad = \quad 0.97 \pm 0.15 \ (stat.)$$

Similarly for the $\sigma(\text{W} + \text{charm})/\sigma(\text{W} + \text{jets})$ ratio:

$$R_c(p_T^{jet} > 20 \text{ GeV}, |\eta^{\text{jet}}| < 2.1) \quad = \quad 0.164 \pm 0.012 \ (stat.)$$

where in this case the charm efficiency corresponding to the region $3 < D_{TCHE} < 20$ is $\epsilon_c = 0.1200 \pm 0.0018 \ (stat.)$.

Both measurements are consistent with the measurements obtained using the SSVHE discriminator, even if no dedicated evaluation of systematic uncertainties for the TCHE method has been performed.

6.6.2 Measurements using a MadGraph W+jets Monte Carlo as reference

The POWHEG W inclusive sample should be a reliable NLO Monte Carlo reference for this analysis, since it is accurate up to W + 1 hard jet plus additional jets in the soft-collinear parton-shower approach. In addition, it contains the relevant $c\bar{c}$ and $b\bar{b}$ contributions from to gluon splitting. However, POWHEG is not expected to reproduce accurately the jet multiplicity and kinematic properties of W + $\geq$ 2 hard jets in the final state, but this limitation has a small impact in the analysis because it is based on the b-tagging of the most significant jet. The jet multiplicity differences between data and MC are also taken into account as systematic uncertainty.

Nevertheless we have repeated the analysis using a (LO) W+jets MadGraph Monte Carlo as reference. The results of the fits to the $\text{W}^+ + charm$ and $\text{W}^- + charm$ distributions are shown in Fig. 6.19. We obtain the following yields:

$$\begin{aligned} N(\text{W}^+ + charm) \quad &= \quad 254.5 \pm 39.5 \ (stat.) \\ N(\text{W}^- + charm) \quad &= \quad 253.4 \pm 36.6 \ (stat.) \end{aligned}$$

Dividing both numbers, we get the following measurement of the W + charm charge ratio:

$$R_c^{\pm}(p_T^{jet} > 20 \text{ GeV}, |\eta^{\text{jet}}| < 2.1) \quad = \quad 1.00 \pm 0.21 \ (stat.)$$

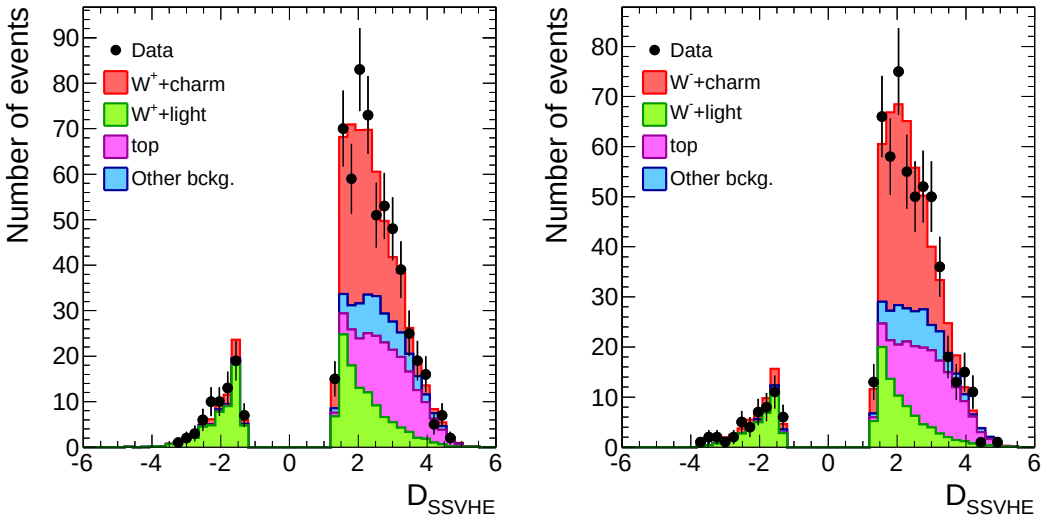

Figure 6.19: Left: fit to the SSVHE discriminator of the $W^+ + charm$ selected sample, using a W+jets MadGraph reference sample to derive. Right: similar fit to the SSVHE discriminator of the $W^- + charm$ selected sample. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution and in blue the sum of the contribution of the other backgrounds. Histograms are stacked.

Similarly for the $\sigma(\mathrm{W}+\mathrm{charm})/\sigma(\mathrm{W}+\mathrm{jets})$ ratio:

$$R_c(p_T^{jet} > 20 \text{ GeV}, |\eta^{\mathrm{jet}}| < 2.1) \quad = \quad 0.129 \pm 0.014 \ (stat.)$$

where the efficiency $\epsilon_c = 0.1320 \pm 0.0018$ $(stat.)$ is found to be consistent - within expected systematic uncertainties quoted before- with the efficiency derived from the POWHEG simulation. The R_c and $R_c^{\pm}$ results thus obtained are consistent with the POWHEG result within the expected uncertainties.

6.7 Final results and comparisons with theory

In summary, we have measured the ratios $R_c^{\pm} \equiv \sigma(\mathrm{W}^+ + charm)/\sigma(\mathrm{W}^- + charm)$ and $R_c \equiv \sigma(\mathrm{W}+\mathrm{charm})/\sigma(\mathrm{W}+\mathrm{jets})$ for leading jets with $p_T > 20$ GeV and $|\eta| < 2.1$ using 36 $\mathrm{pb}^{-1}$of data collected with the CMS detector in 2010. We obtain:

$$\begin{aligned} R_c^{\pm} &= 0.92 \pm 0.20 \ (stat.) \pm 0.04 \ (syst.) \\ R_c &= 0.141 \pm 0.015 \ (stat.) \pm 0.030 \ (syst.) \end{aligned}$$

The ratios are for muons from W decays satisfying $p_T^{\mu} > 25$ GeV and $|\eta^{\mu}| < 2.1$.

The results can be compared with detailed NLO predictions. Values for different PDF sets in the kinematic region $p_T^{jet} > 20$ GeV, $|\eta^{jet}| < 2.1$, $p_T^{\ell} > 25$ GeV, $|\eta^{\ell}| < 2.1$ are shown in Tab. 6.8, where ℓ is the lepton from the W decay. In order to match a realistic experimental scenario and the CMS generator level conditions (W+charm is generated at leading order), the MCFM calculation is performed by combining the charm parton and the potential additional parton from the NLO process into jets using an anti-k_T algorithm (with $R = 0.5$ parameter). To have a theoretical estimate closer to our experimental cuts, all MCFM calculations will use the phase space cuts $p_T^{jet} > 20$ GeV, $|\eta^{jet}| < 2.1$, $p_T^{\mu} > 25$ GeV, $|\eta^{\mu}| < 2.1$.

The theory uncertainties include just the PDF contributions (at 68% CL) and have been calculated following the established recipes for each set of PDFs. CT10 (93) predictions for $R_c^{\pm}$ are larger than the ones from MSTW08 (94) or NNPDF21 (95), beyond the expected uncertainties (at 68% CL). This is most likely due to the CT10 assumption of equal PDFs for strange and anti-strange quarks. MSTW08 and NNPDF21 do not assume this constraint and use experimental data (from NuTeV and others) to constrain both PDFs independently. Concerning R_c, MSTW08 and NNPDF21 predictions have much smaller uncertainties than CT10. CT10 has more relaxed criteria for the overall strange content of the proton, while MSTW08 strongly constrains the strange PDF shapes to follow anti-u and anti-d shapes (thus constraining R_c in practice). NNPDF21 has many more experimental input measurements to constrain strange PDFs and also employs a positiveness constraint that reduces the uncertainty further (also in a region where no data are available, $x < 10^{-2}$).

The results are in agreement with the theoretical prediction given in Tab. 6.8. Fig. 6.20 shows a comparison of our data measurements with those predictions.

Ratio	MCFM (CT10)	MCFM (MSTW08)	MCFM (NNPDF21)
$R_c^{\pm}$	0.993 ± 0.008	0.926 ± 0.028	0.950 ± 0.007
R_c	0.125 ± 0.017	0.115 ± 0.002	0.104 ± 0.005

Table 6.8: $R_c^{\pm}$ and R_c predictions from MCFM at NLO. Kinematic cuts are: $p_T^{jet} > 20$ GeV, $|\eta^{jet}| < 2.1$, $p_T^{\ell} > 25$ GeV, $|\eta^{\ell}| < 2.1$. Partons are joined using an anti-k_T algorithm with $R = 0.5$ parameter. The quoted values correspond to different PDF choices and only PDF variations (at 68% CL) are considered for the total uncertainties.

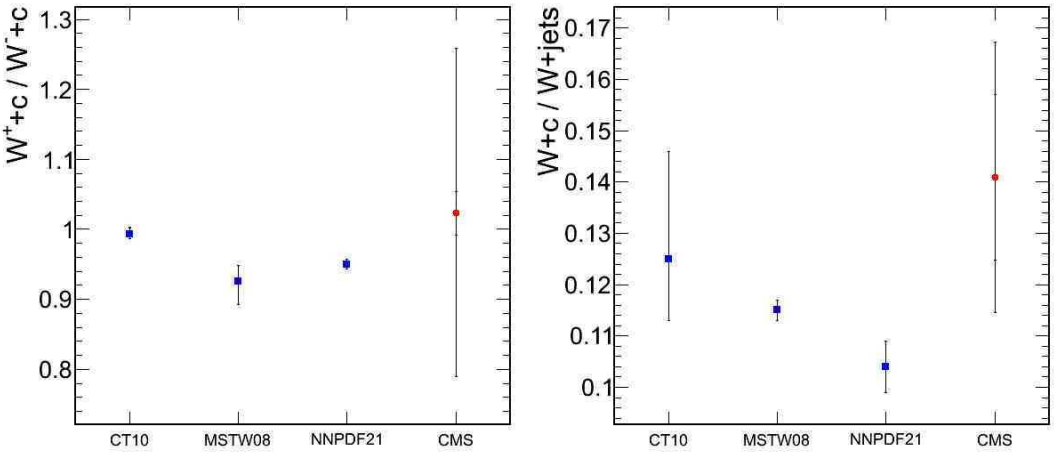

Figure 6.20: Comparison of the CMS measured values for $R_c^{\pm}$ and R_c with MCFM NLO predictions for different PDF sets (CT10, MSTW08, NNPDF21). Only PDF theoretical uncertainties within each PDF set are shown.

As already advanced the discrepancy among the different predictions exceed their uncertainties and a precise experimental measurement would help to clear the situation.

6.8 Prospects for future improvements

These measurements are a first step towards a precise understanding of the strange and anti-strange parton distribution functions of the proton at the electroweak scale probed by the LHC. With more data incoming, this analysis opens new oportunities to improve our knowledge on the strangeness composition in the proton.

With 5 fb^{-1} recorded in 2011 run the amount of data is already large enough to place strong constraints on the strange composition in the proton. A simulation study performed on the differential measurement of the R_c parameter with jet pseudorapidity with this amount of data indicates that it already gives valuable information on the strange composition. In Fig. 6.21 we can appreciate the improvement on the strange quark component of the NNPDF PDF set when the information coming from a W+c measurement at LHC, with 5 fb^{-1} data is included. The information provided by this extended analysis would not only increase our knowledge on the strange composition, but also would allow the NNPDF group to use only collision data (from HERA,

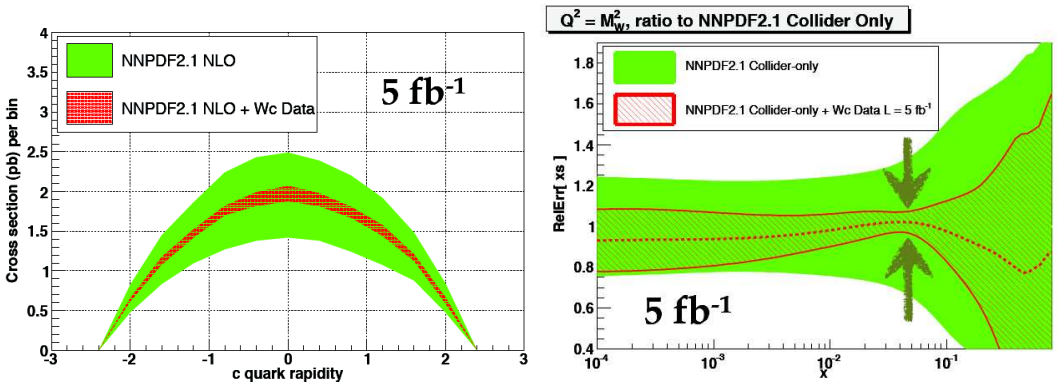

Figure 6.21: Left: Improvement on the strange quark component for different eta bins for the NNPDF PDF set with and without data coming from LHC W+c study. Right: Improvement on the strange quark densities in the proton for different x parameter values, with current NNPDF PDF set and the new one (including the results from a W+c analysis with 5 fb^{-1} LHC data) (82).

Tevatron and now LHC) to perform the fits. This latter advantage would be a major improvement since older data coming from neutrino experiments would not be necessary anymore. Including neutrino experiment data entails the modeling of nuclear corrections for deuterium or heavy nuclei and bigger pQCD uncertainties from higher scales are involved.

The improvements proposed previously imply a challenging analysis in the W+c channel. The simulation shown in Fig. 6.21 requires the total experimental error on the R_c parameter to be of the order of 5%. In the analysis presented in this Chapter on the R_c parameter the error is estimated to be of the order of 20%, way too far from the precision required by the PDF groups. With a bigger amount of data, new techniques can be used to reduce the systematic errors associated to the measurement. In particular, the tracking error and the top template, which amount each to 14%, can be reduced.

The top template error can be decreased by reducing the top related backgrounds to the minimum. More stringent requirements on the number of jets in the event (i.e. setting to only one the number of jets with energy bigger than 40 GeV allowed in the event) would strongly reduce the amount of top background in the sample. By imposing a more restrictive requirement on the number of jets we reduce the $t\bar{t}$ background by a 75%, 30% the single-top background and only 5% the signal. MadGraph MC should

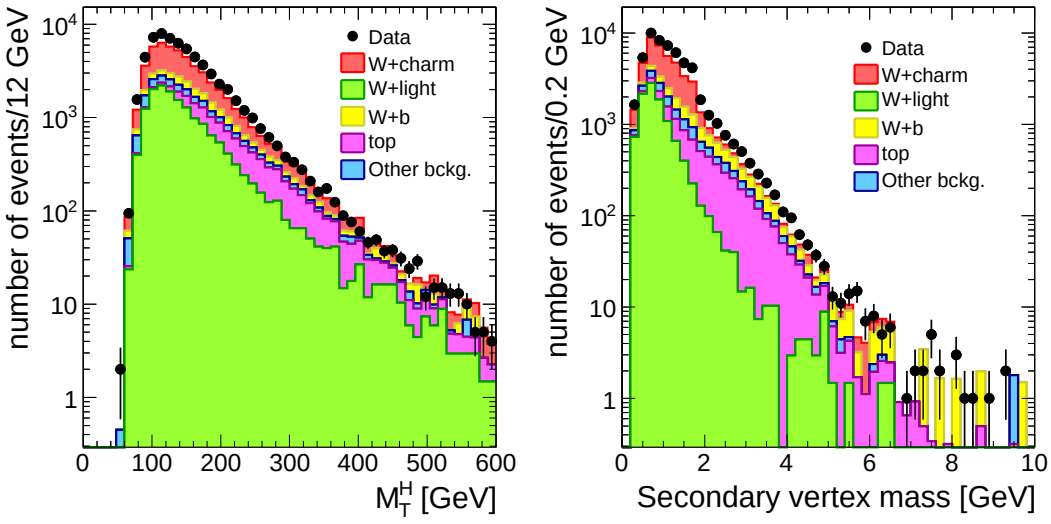

Figure 6.22: M_T^H distribution (left) and secondary vertex mass distribution (right) for the 2011 dataset. Thresholds at 300 GeV and 2 GeV (respectively) are suggested to remove the top related background. In red the W+c signal, in green the W+light quark contribution, in purple the top contribution, in yellow the W+b and in blue the sum of the contribution of the other backgrounds. Given the higher relevance of the W+b contribution, it is now presented explicitly. Histograms are stacked.

be used in such case, since POWHEG does not correctly describe events with high jet multiplicity. To further reduce the top related backgrounds, and profiting from the high statistics available, stringent requirements on the invariant mass of the secondary vertex or in the transverse mass of the jets plus W system (hereafter M_T^H) can be set. In Fig. 6.22 these distributions are shown for 2.2 fb^{-1} of 2011 data. Requiring the secondary vertex mass to be smaller than 2 GeV and the M_T^H smaller than 300 GeV removes partially the top background and consequently the systematic error associated to it. These last requirements reduce $t\bar{t}$ background by an additional 35%, the single-top background by 23% and the signal by a 1%. With these new requirements, the top related background is not anymore the dominant background, amounting to some 3% of the total data sample, being the dominant backgrounds W+light jets (30%) and W+b jets (7%).

To reduce the tracking related systematic error the tagging efficiency should be evaluated directly from the data. The proposed method, already checked successfully

on 2011 data, consists in computing this efficiency in a sample with dimuon events. W+jet dimuon events are those events with a second muon coming from the decay of a charm hadron in the jet. Signal events in this sample (as can be seen in Fig. 6.1) are characterized by having two muons (one from the W boson, the other from the c quark jet) with different sign (opposite sign, OS, muons).

With high statistics we can afford to select two subsamples: events having a secondary vertex characteristic of the decay of a heavy flavour quark (or c-quark) and events with two muons of opposite charge in the final state, one of them characteristic of the semileptonic decay of the heavy flavour quark (or c-quark). In fact we are using an independent tagging method (a lepton tag). The normalization of these two samples would give us information about the normalization scale of the signal and the efficiency for each of the sets, i.e. the signal normalization and the c-tagging efficiency from the SSVHE set and the signal normalization and the c-lepton tagging from the other set. The intersection of these two sets (those events b-tagged and with two muons of opposite sign) would give information of the signal normalization, the b-tagging efficiency and the lepton tagging efficiency. The system:

$$N^{tagged}_{SSVHE} = S \cdot \epsilon_V$$
$$N^{tagged}_{lepton} = S \cdot \epsilon_l$$
$$N^{tagged}_{SSVHE+lepton} = S \cdot \epsilon_V \cdot \epsilon_l$$

can be solved by means of a set of discriminant distributions from the three plots for each of the 3 subsamples (where S is the signal normalization, ϵ_V the vertex efficiency and ϵ_l the lepton efficiency). To reinforce the W+c-W+b separation, we can use a variable with high discriminant power for the dimuon subsample and the intersection of it with the b-tagged one. The variable chosen is the $p_{T,rel}$, defined as the p_T of the muon relative to the jet direction. b quark jets are less collimated, and a higher value on this $p_{T,rel}$ variable is expected for this contribution. The $p_{T,rel}$ distribution for the OS dimuon events subsample and the intersection b-tag dimuon OS sample is shown Fig. 6.23.

We can further exploit the possibilities of this large recorded statistics and use the set of same sign dimuon events to control the light background (see Fig. 6.24). The fit is then performed over 5 histograms for each W charge: the SSVHE discriminant for the b-tagged events and the $p_{T,rel}$ distribution for the dimuon tagged events and for

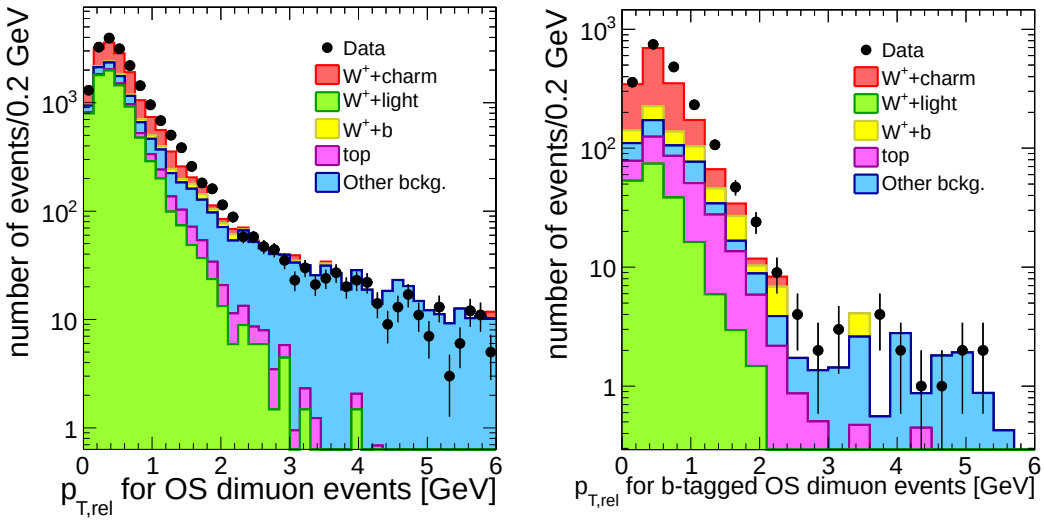

Figure 6.23: $p_{T,rel}$ distribution for opposite sign dimuon events for the lepton tagged set (left) and for the vertex and lepton tagged set (right). In red the W+c signal, in green the W+light quark contribution, in purple the top contribution, in yellow the W+b and in blue the sum of the contribution of the other backgrounds. Given the higher relevance of the W+b contribution, it is now presented explicitly. Histograms are stacked.

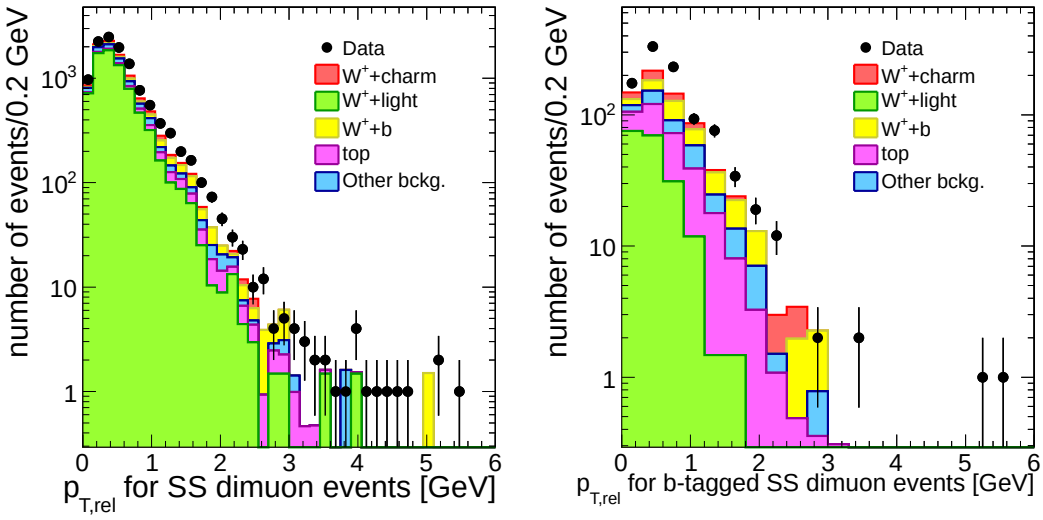

Figure 6.24: $p_{T,rel}$ distribution for same sign dimuon events for the lepton tagged set (left) and for the vertex and lepton tagged set (right). In red the W+c signal, in green the W+light quark contribution, in purple the top contribution, in yellow the W+b and in blue the sum of the contribution of the other backgrounds. Given the higher relevance of the W+b contribution, it is now presented explicitly. Histograms are stacked.

the b-tagged dimuon events. A summary of the subsets used in the analysis and the information given by each set is presented in Tab. 6.9.

The method is well established and preliminary results are encouraging. Reaching the desired precision seems to be feasible.

	Positive Discriminator	Negative Discriminator
B-tagged events	D_{SSVHE} Signal Norm., c-tagging eff.	D_{SSVHE} background control
	OS	SS
Dimuon events	$p_{T,rel}$ Signal Norm., dimuon eff.	$p_{T,rel}$ background control
Dimuon-b-tagged events	$p_{T,rel}$ Signal Norm., c-tagging eff, dimuon eff.	$p_{T,rel}$ background control

Table 6.9: Subsamples used in the proposed analysis and the information given in each set.

6. MEASUREMENT OF ASSOCIATED PRODUCTION OF W BOSONS WITH CHARMED JETS

7

Conclusions

This thesis presents several analysis involving muons detected in the CMS experiment at the LHC. The final goal was the study of the inclusive W boson production and the associated production of W bosons with charm-jets in pp collisions at $\sqrt{s} = 7$ TeV with the data collected by CMS during the 2010 LHC run. Being the muons fundamental elements of many analysis and in particular of the ones presented in this report, the first part of the thesis is dedicated to performance studies of the muon detector, in particular of the central detector (DT system) with cosmic data, in the context of the final commissioning of the detector already in place in its final location.

The central muon system is based on the measurement of the drift time of the electron cloud produced by the passage of a muon through the detector. One of the main quantities to be calibrated and monitored regularly to ensure a proper translation of the physical quantity measured, drift time, to muon track position is the drift velocity. The drift velocity in the chambers was measured and found to be homogeneous in all of them once the uncertainty in the timing produced by the asynchronous arrival of the cosmic muons is corrected for. It was also checked that the measured velocity varies with the angle of incidence of muons. There was up to a 3% increase in the apparent drift velocity for muons with an incidence angle of 30^0, the maximum angle expected for muons produced at the interaction vertex. Finally, it was confirmed that the drift velocity is not affected by the residual magnetic field present in the chambers, except in the innermost stations of the outermost wheels, where the remnant radial magnetic field was largest. The maximal variation observed is $\sim 2.6\%$.

One of the design objectives of the muon barrel spectrometer was to guarantee a 100% muon detection and reconstruction efficiency. This is achieved with a highly redundant design, with several layers of detector units, complemented with an optimized software to perform the track reconstruction from the hits recorded in the chambers. The muon reconstruction efficiency in the DT system was evaluated with a new method proposed in this thesis. A careful study taking into account the geometry of the detector, dead zones, overlaps, the irregular illumination of cosmic muons in the different stations, the different momentum spectra of muons reaching top and bottom parts of the detector wheels etc. was necessary in order to obtain an unbiased result. The reconstruction efficiency was shown to be very high, larger than 99% in the bending plane and bigger than 95% in the logitudinal direction.

An accurate knowledge of the muon performance in terms of momentum assignment (scale and resolution) is crucial in any physics analysis. Should any effect potentially affecting the muon measurement had been found, proper corrections would have had to be incorporated in order to ensure a final result free of any bias. A novel approach has been proposed in this work to determine the muon momentum scale and resolution. The description of the behaviour of the CMS detector and of the software algorithms implemented for muon reconstruction is taken as the baseline description, and new terms modifying it are included in the MC to reproduce the observed characteristics of the data. These additional terms are evaluated by careful comparison of the MC prediction of the reconstructed dimuon mass distribution in the Z mass region with the experimental distribution.

The method has been applied to the data collected by CMS during the 2010 LHC run. It showed that the muon momentum scale and resolution are very well described in the MC and only residual effects were found. A periodical dependence on the azimuthal angle ϕ was found, with the same amplitude but opposite sign for positive and negative muons. A certain modulation of the effect as a function of the muon pseudorapidity was also observed. Overall the effect is very small and it was not considered to be large enough that a correction is needed for the physics analysis reported in this thesis. These effects were properly quantified and parametrized and delivered to the physics groups of the collaboration working on analysis relying on muon measurement.

The inclusive production of W boson has been studied. The total production cross sections as well as the production cross section for positive and negative W and their

ratio were measured to be:

$$\begin{aligned}
\sigma(\mathrm{W}) &= 10.10 \pm 0.03\ (stat.) \pm 0.16\ (syst.) \pm 0.40\ (lumi.) \\
\sigma(\mathrm{W}^+) &= 5.93 \pm 0.02\ (stat.) \pm 0.10\ (syst.) \pm 0.24\ (lumi.) \\
\sigma(\mathrm{W}^-) &= 4.17 \pm 0.02\ (stat.) \pm 0.09\ (syst.) \pm 0.17\ (lumi.) \\
\sigma(\mathrm{W}^+)/\sigma(\mathrm{W}^-) &= 1.421 \pm 0.008\ (stat.) \pm 0.11\ (syst.)
\end{aligned}$$

The result is in agreement with and supersedes the first CMS measurement, performed with a smaller data sample. The selection process was carefully optimized aiming for a reduction of the main experimental systematic uncertainty, arising from the background subtraction, keeping at the same time under control the uncertainty due to the theoretical uncertainties in the acceptance correction factors. The measurement also profited from the larger data sample that allowed a better understanding of the data and a reduction of certain experimental uncertainties such as those associated to lepton efficiency, signal modeling etc. The precision achieved in the final result is at the level of 1.6% in the total cross section, 1.7% and 2.1% for the individual W^+ and W^- cross sections (excluding the luminosity uncertainty) and 2.5% for the ratio. The experimental contribution to these uncertainties is small, being the theoretical uncertainties the major contributors to the final budget. These results have validated next-to-next-to-leading order QCD cross section calculations based on recent parton distribution functions.

Finally, the large data sample collected permitted the access to W final states with a rather small production cross section. The process studied was the associated production of a W boson with a c-quark jet. This final state is produced dominantly by the interaction of a strange quark and a charm quark from the sea, thus directly probing the strange quark content of the proton in a wide range on the x-Bjorken variable. For that purpose the CMS algorithms for heavy quark tagging (mainly b-tagging) were successfully employed for c-tagging for a first time in the experiment. The measurement was formulated in terms of cross section ratios of the production of a W boson accompanied by a charm jet over the total associated production W+jets. A second measurement was the charged ratio production of W+charm, $(\mathrm{W}^+ + \mathrm{charm})/(\mathrm{W}^- + \mathrm{charm})$. The

measured ratios were:

$$\begin{aligned} R_c^{\pm} &= 0.92 \pm 0.20\ (stat.) \pm 0.04\ (syst.) \\ R_c &= 0.141 \pm 0.015\ (stat.) \pm 0.030\ (syst.) \end{aligned}$$

The $R_c^{\pm}$ ratio is limited by the statistical precision ($\sim 20\%$), the systematic uncertainty having a reduced impact ($\sim 4\%$). The measured R_c ratio, instead, is affected by a rather large systematical uncertainty ($\sim 20\%$), arising mainly from the c-tagging efficiency and top background subtraction.

These measurements constitute a first step towards a comprehensive understanding of the W+charm production in pp collisions at high energy. Based on the experience acquired with the present study, the analysis strategy to be carried with the data collected in 2011 is already deployed. Projections from the current measurements indicate the feasibility to significantly reduce the experimental uncertainty thus contributing to an improvement in the determination of the strange quark Parton Distribution Function.

References

[1] M. B. *et al.*, "The LHC design report v.3: the LHC Infrastructure and General Services," *CERN-2004-003-V-3*, 2004. 5, 11

[2] O. B. *et al.*, "The LHC design report v.1: the LHC Main Ring," *CERN-2004-003-V-1*, 2004. 5, 11, 13

[3] O. B. *et al.*, "The LHC design report v.2 : the LHC Infrastructure and General Services," *CERN-2004-003-V-2*, 2004. 5, 11

[4] M. Herrero, "The standard model," *http::/arxiv.org/abs/hep-ph/9812242v1*, 2007. 5

[5] C. Amsler *et al.*, "Review of particle physics," *Phys. Lett.*, vol. B667, 2008. 6, 29, 57

[6] LEP Electroweak Working Group, ATLAS Collaboration, http://lepewwg.web.cern.ch/LEPEWWG/. 7

[7] R. Barate *et al.*, "Search for the standard model Higgs boson at LEP," *Phys. Lett.*, vol. B565, pp. 61–75, 2003. 7

[8] Tevatron New Phenomena and Higgs Working Group, "Combined CDF and DØ Upper Limits on Standard Model Higgs-Boson Production with up to 6.7 fb^{-1} of Data," *webpage, arXiv:1007.4587 [hep-ex]*, 2010. 7

[9] "Combined Standard Model Higgs boson searches with up to 2.3 fb^{-1} of pp collisions at $\sqrt{s} = 7$ TeV at the LHC," Tech. Rep. ATLAS-CONF-2011-157, CERN, Geneva, Nov 2011. 7

[10] S. P. Martin, "A supersymmetry primer," *arXiv:hep-ph/9709356v6*, 2011. 10

[11] K. R. Dienes, "String theory and the path to unification: A review of recent developments," *arXiv:hep-th/9602045v3*, 1997. 10

[12] C. T. Hill and E. H. Simmons, "Strong dynamics and electroweak symmetry breaking," *arXiv:hep-ph/0203079v3*, 2002. 10

[13] M. J. Hadley, "Classical dark matter," *arXiv:gr-qc/0701100v1*, 2007. 10

[14] I. Bigi, "CP violation - an essential mystery in nature's grand design," *arXiv:hep-ph/9712475v1*, 1997. 10

[15] M. Awramik, M. Czakon, A. Freitas, and G. Weiglein, "Precise prediction for the W boson mass in the Standard Model," *Phys. Rev. D*, vol. 69, p. 053006, Mar 2004. 11

[16] **CDF** Collaboration, "Observation of top quark production in $p\overline{p}$ collisions with the collider detector at Fermilab," *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, Apr 1995. 11

[17] **ALICE** Collaboration, "ALICE: Technical proposal for a large ion collider experiment at the CERN LHC," *CERN-LHCC-95-71*, 1995. 14

[18] **ATLAS** Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *JINST 3 S08003*, 2008. 14

[19] **CMS** Collaboration, "The CMS experiment at the CERN LHC," *JINST 0803:S08004*, 2008. 14, 87

[20] **LHCb** Collaboration, "LHCb technical design report: Reoptimized detector design and performance," *CERN-LHCC-2003-030*, 2003. 14

[21] **CMS** Collaboration, "CMS Physics Technical Design Report Volume 1: Detector Performance and Software," *CERN-LHCC-2006-001, http://cdsweb.cern.ch/record/922757*. 15, 20, 23

[22] G. Bayatian *et al.*, "CMS physics Technical Design Report, Volume II: Physics performance," *J.Phys.G*, vol. G34, pp. 995–1579, 2007. 15

[23] **CMS** Collaboration, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays," *JINST*, vol. 5, p. T03021, 2010. 17

[24] **CMS** Collaboration, "The CMS tracker system project : Technical Design Report," *CMS-TDR*, no. CERN-LHCC-98-006; CMS-TDR-005, 1997. 18

[25] **CMS** Collaboration, "The CMS tracker: addendum to the Technical Design Report," *CMS-TDR*, no. CERN-LHCC-2000-016; CMS-TDR-005-add-1, 2000. 18

[26] **CMS** Collaboration, "Measurement of Momentum Scale and Resolution using Low-mass Resonances and Cosmic Ray Muons," *CMS-PAS-TRK-10-004*, 2010. 21

[27] **CMS** Collaboration, "Tracking and Vertexing Results from First Collisions," *CMS-PAS-TRK-10-001*, 2010. 22

[28] **CMS** Collaboration, "The CMS electromagnetic calorimeter project : Technical Design Report," *CMS-TDR*, no. CERN-LHCC-97-033; CMS-TDR-004, 1997. 20

[29] **CMS** Collaboration, "Electromagnetic calorimeter commissioning and first results with 7 TeV data," *CMS-PAS-EGM-10-002*, 2010. 24, 25

[30] **CMS** Collaboration, "The CMS hadron calorimeter project : Technical Design Report," *CMS-TDR*, no. CERN-LHCC-97-031; CMS-TDR-002, 1997. 23

[31] **CMS** Collaboration, "Jet Performance in pp Collisions at 7 TeV," *CMS-PAS-JME-10-003*, 2010. 26

[32] **CMS** Collaboration, "Jet energy resolution in CMS at $\sqrt{s} = 7$ TeV," *CMS note JME-10-014*, 2011. 28

[33] **CMS** Collaboration, "The CMS muon project : Technical Design Report," *CMS-TDR*, no. CERN-LHCC-97-032; CMS-TDR-003, 1997. 29, 31

[34] P. Igo-Kemenes, "Searches for higgs bosons." pdg.lbl.gov/2006/reviews/higgs_s055.pdf. 30

[35] **CMS** Collaboration, "CMS physics technical design report, volume II: Physics performance," *J. Phys. G: Nucl. Part. Phys*, 2007. 30

[36] G. Abbiendi *et al.*, "Muon reconstruction in the CMS detector," *AN-2008/097*, 2008. 37, 60, 87, 143

[37] **CMS** Collaboration, "Performance of the CMS drift tube chambers with cosmic rays," *JINST*, vol. 5, no. 03, p. T03015, 2010. 39, 41, 54

[38] C. Grupen, *Particle Detectors*. Cambridge University Press, 1996. 88

[39] S. Agostinelli *et al.*, "Geant4: a simulation toolkit," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 506, no. 3, pp. 250 – 303, 2003. 90

[40] **CMS** Collaboration, "Tracking and vertexing results from first collisions," *CMS PAS TRK-10-001 (2010)*. 92

[41] **CMS** Collaboration, "Tracking and vertexing results from first collisions," *CMS PAS TRK-10-004 (2010)*. 92

[42] P. Nason, "A new method for combining NLO QCD with shower Monte Carlo algorithms," *JHEP*. 95, 108, 124, 142, 178

[43] P. S.Frixione and C.Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method," *JHEP*. 95, 142

[44] J.Campbell and K.Ellis, "A Monte Carlo for FeMtobarn processes at Hadron Colliders." `http://mcfm.fnal.gov/`. 96, 177

[45] A. D. Martin *et al.*, "Parton distributions for the LHC," *arXiv:0901.0002v3 [hep-ph]*. 96, 124

[46] **CMS** Collaboration, "Results on local muon reconstruction in DT chambers from analysis of CRAFT data," *CMS CFT 09-012*. 95

[47] R. Castello, *Alignment of CMS Tracker detector using cosmic ray particles and its impact on early physics performance*. PhD thesis, Turin University, 2010. 109

[48] M. Borgia, *The CMS Silicon Tracker: from the performance study in cosmic runs to the measurement of the momentum resolution from the Z line-shape in early data*. PhD thesis, Turin University, 2010. 121

[49] T. Sjostrand *et al.*, "Pythia 6.2: Physics and manual," *arXiv:hep-ph/0108264v1*. 122, 142

[50] D. Bourilkov *et al.*, "LHAPDF: PDF use from the Tevatron to the LHC," *arXiv:hep-ph/0605240v2*. 124

[51] J. Pumplin *et al.*, "New generation of parton distributions with uncertainties from global QCD analysis," *arXiv:hep-ph/0201195v3*. 124

[52] C.Balazs *et al.*, "RESBOS package. QT resummation portal." http://hep.pa.msu.edu/resum. 125

[53] P. Langacker, *Precision test of the standard electroweak model*. World scientific publishing, 1995. 129

[54] U. Baur *et al.*, "The charm content of W+1 jet events as a probe of the strange quark distribution function.," *arXiv:hep-ph/9308370v1*, 1993. 129

[55] **CDF** Collaboration, "First run II measurement of the W boson mass at the Fermilab Tevatron," *Phys. Rev. D*, vol. 77, p. 112001, Jun 2008. 130

[56] **CMS** Collaboration, "Search for W' in the leptonic channels in pp collisions at $\sqrt{s}$ =7 TeV," *CMS-EXO-11-024*, 2011. 131

[57] **CMS** Collaboration, "Measurements of inclusive W and Z cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the CMS experiment," *JHEP*, vol. 2011, pp. 1–40, 2011. 10.1007/JHEP01(2011)080. 130, 140, 145, 171

[58] "Measurement of the lepton charge asymmetry in inclusive W production in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 2011, pp. 1–31, 2011. 10.1007/JHEP04(2011)050. 130

[59] C. F. Berger *et al.*, "Next-to-Leading Order QCD Predictions for W+3-Jet Distributions at Hadron Colliders," *Phys. Rev.*, vol. D80, p. 074036, 2009. 131

[60] **CMS** Collaboration, "Measurement of the Drell-Yan cross section in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 2011, pp. 1–41, 2011. 10.1007/JHEP10(2011)007. 131

[61] **CMS** Collaboration, "Measurement of the polarization of W bosons with large transverse momenta in W + jets events at the LHC," *Phys. Rev. Lett.*, vol. 107, p. 021802, Jul 2011. 131

[62] **CMS** Collaboration, "Jet production rates in association with W and Z bosons in pp collisions at $\sqrt{s} = 7$ TeV," *arXiv:1110.3226*, 2011. 132

[63] K. Melnikov and F. Petriello, "W boson production cross section at the Large Hadron Collider with $\mathcal{O}(\alpha_s^2)$ corrections," *Phys. Rev. Lett.*, vol. 96, p. 231803, Jun 2006. 132, 135

[64] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre, "Precise predictions for W+4-jet production at the Large Hadron Collider," *Phys. Rev. Lett.*, vol. 106, p. 092001, Mar 2011. 132

[65] R. K. Ellis, K. Melnikov, and G. Zanderighi, "W+3 jet production at the Tevatron," *Phys. Rev.*, vol. D80, p. 094002, 2009. 132

[66] C. F. Berger, Z. Bern, L. J. Dixon, F. Febres Cordero, D. Forde, T. Gleisberg, H. Ita, D. A. Kosower, and D. Maitre, "Next-to-leading order QCD predictions for Z,γ^* + 3-jet distributions at the Tevatron," *Phys. Rev. D*, vol. 82, p. 074002, Oct 2010. 132

[67] **CMS** Collaboration, "Measurement of associated charm production in W final states at $\sqrt{s} = 7$ TeV," *CMS note EWK-11-013*, 2011. 132

[68] **CMS** Collaboration, "Measurement of the Z/γ b-jet cross section in pp collisions at $\sqrt{s} = 7$ TeV," *CMS note EWK-10-012*, 2011. 132

[69] "Measurement of the inclusive Z cross section via decays to tau pairs in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 2011, pp. 1–32, 2011. 10.1007/JHEP08(2011)117. 132

[70] **CMS** Collaboration, "Observation of $W \to \tau\nu$ production in pp collisions at $\sqrt{s}$ = 7 TeV," *CMS-PAS-EWK-11-002*, 2011. 132

[71] **CMS** Collaboration, "Measurement of forward-backward asymmetry of lepton pairs and the weak-mixing angle at CMS," *CMS note EWK-10-011*, 2011. 133

[72] **CMS** Collaboration, "Measurement of wγ and zγ production in pp collisions at sqrt(s) = 7 tev," *Physics Letters B*, vol. 701, no. 5, pp. 535 – 555, 2011. 133

[73] **CMS** Collaboration, "Measurement of w^+w^- production and search for the higgs boson in pp collisions at sqrt(s) = 7 tev," *Physics Letter B*, vol. 699, no. 1-2, pp. 25 – 47, 2011. 133

[74] **CMS** Collaboration, "Measurement of the weak mixing angle with the Drell-Yan process in proton-proton collisions at the LHC," *Phys. Rev. D*, vol. 84, p. 112002, Dec 2011. 133

[75] A. Straessner, *Electroweak physics at LEP and LHC*. Springer, 2010. 134

[76] K. Melnikov and F. Petriello, "Electroweak gauge boson production at hadron colliders through $\mathcal{O}(\alpha_s^2)$," *Phys. Rev. D*, vol. 74, p. 114017, Dec 2006. 135

[77] C. Diez *et al.*, "Establishing the method to measure the single muon efficiency from $Z \to \mu\mu$ events with the first data," *CMS AN-2010/233*. 151

[78] C. Diez, *Searches for New Heavy W' Bosons Decaying into Muon and Neutrino in pp Collisions at 7 TeV in the CMS Experiment at the LHC*. PhD thesis, Universidad Complutense, 2011. 151

[79] M. Cepeda, *Measurement of the inclusive production cross-section of W bosons in proton-proton collisions at 7 TeV in the CMS experiment at the LHC*. PhD thesis, Universidad Complutense de Madrid, 2011. 153

[80] **CMS** Collaboration, "Measurement of the inclusive W and Z production cross sections in pp collisions at $\sqrt{s} = 7$ TeV with the CMS experiment," *JHEP*, vol. 2011, pp. 1–76, 2011. 10.1007/JHEP10(2011)132. 169

[81] **ATLAS** Collaboration, "Measurement of the inclusive $W^{\pm}$ and Z/γ cross sections in the electron and muon decay channels in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector," 2011. 172

[82] J.Rojo, "W+c data with aMCatNLO an their possible impact on strangeness." Presentation at the Working Group on Electroweak precision measurements at the LHC, November 2011. 175, 206

[83] M. W. Krasny, F. Dydak, F. Fayette, W. Placzek, and A. Siodmok, "$\Delta M_W < 10$ MeV/c^2 at the LHC: a forlorn hope?," *Eur. Phys. J.*, vol. C69, pp. 379–397, 2010. 175

[84] J. Alwall *et al.*, "MadGraph/MadEvent v4: The New Web Generation," *JHEP*, vol. 09, p. 028, 2007. 179

[85] **CMS** Collaboration, "Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV," *arXiv:1107.0330v1*, 2011. 179

[86] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP*, vol. 2008, no. 04, p. 063, 2008. 181

[87] **CMS** Collaboration, "Particle-flow event reconstruction in CMS and performance for jets, taus and MET," *CMS PAS PFT-09-001*, 2009. 181

[88] **CMS** Collaboration, "Commissioning of the particle-flow reconstruction in minimum-bias and jet events from pp collisions at 7 TeV," *CMS PAS PFT-10-002*, 2010. 181

[89] **CMS** Collaboration, "Commissioning of b-jet identification with pp collisions at $\sqrt{s} = 7$ TeV," *CMS PAS*, vol. BTV-10-001, 2010. 186, 200

[90] **CMS** Collaboration, "Jet energy corrections determination at 7 TeV," *CMS PAS*, vol. JMT-10-010, 2010. 195

[91] **CMS** Collaboration, "Jet performance in pp collisions at 7 TeV," *CMS PAS*, vol. JMT-10-003, 2010. 195

[92] **CMS** Collaboration, "Performance of b-jet identification with pp collision at $\sqrt{s} = 7$ TeV," *CMS PAS BTV-11-001*, 2011. 197

[93] H. Lai *et al.*, "CT10 parton density functions," *arXiv:1007.2241*. 198, 204

[94] A. Martin *et al.*, "MSTW2008 parton density functions," *arXiv:0901.0002*. 198, 204

[95] R. Ball *et al.*, "NNPDF2.1 parton density functions," *arXiv:1101.1300*. 198, 204

[96] L. Gladilin, "Charm hadron production fractions," *arXiv:hep-ex/9912064v1*. 198