SHARIF UNIVERSITY OF TECHNOLOGY
PHYSICS DEPARTMENT

Ph.D. Thesis

Search for new physics in same-sign dilepton channel using 35 pb^{-1} of the CMS data

By:
Hamed BAKHSHIANSOHI

Supervisor:
Professor Farhad ARDALAN

Co-Supervisor:
Professor Luc PAPE

October 2011

Contents

List of Figures

List of Tables

Acknowledgement

This research project would not have been possible without the support of many people.

I would like to acknowledge the Institute for Research in Fundamental Sciences (IPM) and in particular, the School of Particles and Accelerators that gave me the opportunity to perform my research at CERN. I would like to thank professor Hessamaddin Arfaei, the IPM team leader in CERN program, whose support has been a big source of motivation for me during my PhD.

It is difficult to express my deep appreciation to my supervisor, professor Farhad Ardalan for his invaluable advices in my scientific life. His inspiration and enthusiasm helped me to find my interests in the exciting world of high energy physics.

Along with my research, I had constructive and fruitful discussions with my co-supervisor, professor Luc Pape, who generously and patiently devoted his time for me to acquire expertise in experimental particle physics.

I would like to thank researchers in the School of Particles and Accelerators, Dr. Saeid Paktinat, Dr. Mojtaba Mohammadi, Dr. Mohsen Khakzad, Dr. Majid Hashemi and Dr. Batool Safarzadeh as well as my friends, Ali, Abdollah, Maryam and Nadjieh. Thank you all for making our IPM-CMS meetings so useful and productive.

The chance to collaborate with Dr. Filip Moortgat, Benjamin Steiger and Predrag Milenovic had a large impact on my PhD progress. I would like to thank them, all.

Introduction

The Standard Model of particle physics has been very successful in describing the physics of subatomic particles and their interactions. The discovery of W and Z bosons as weak force carriers in 1982 together with the observation of the top quark as the heaviest quark in 1995 are among its notable achievements. The model predicts the existence of the Higgs boson as the responsible for the electroweak symmetry breaking. This not yet discovered particle plays also the crucial role of giving mass to the other particles in the theory.

Despite all of its successes, the Standard Model has several shortcomings and theoretical inconsistencies. This has led to many extensions to the model, amongst which Supersymmetry (SUSY) is one of the most favorables. The new symmetry between particles, proposed by SUSY, results in solutions for many of the questions that have no clear answer in the Standard Model framework.

In the quest for the Higgs boson and to search for the new physics, the Large Hadron Collider (LHC) is designed and constructed in a 27 km long tunnel at CERN. The LHC collides two intense proton beams with a center of mass energy of 7 TeV. The Compact Muon Solenoid (CMS) is one of the four detectors present at four collision points on the LHC ring.

Equipped with a strong magnetic field and excellent subdetectors, the CMS detector is able to record collision events and to reconstruct different physics objects including leptons with a very good precision.

One of the fascinating signatures of new physics is events with two same-sign leptons in the final state. There are limited sources for the same-sign dilepton events in the Standard Model. Hence backgrounds of this channel are under control.

The same-sign dilepton events in SUSY are produced via the decay of two squarks into two same-sign charginos where the leptonic decay of charginos occurs via sleptons or W bosons. Thanks to the particular kinematics, the mass spectrum of contributing particles in this process can be extracted which is a great advantage of the channel in case of discovery.

In the first chapter of this thesis, the theory of Supersymmetry is described after a short review of the Standard Model. It is followed by an overview of the minimal supersymmetric extension to the Standard Model (MSSM) and its phenomenology. A description of the LHC machine and the CMS detector together with their performances

during 2010 data taking is given in the second and the third chapters. The fourth chapter is devoted to the algorithms used for the reconstruction of the physics objects such as muons, electrons and jets. As the charge of the electron plays an important role in our analysis, different methods for the electron charge measurement are discussed in the same chapter where a new method based on the shape of calorimeter hits is also developed.

The fifth chapter is dedicated to the search for new physics in the same-sign dilepton channel using the full data set collected by CMS in 2010 data taking. Besides the appropriate event selection, the effect of the electron charge mis-identification on background contamination is investigated in details. A method to estimate backgrounds due to fake leptons is developed and used in this analysis. At the end, the event yield in data is compared to the background expectation and new bounds on new physics are set.

1 Standard Model and Supersymmetry

Our current knowledge about the physics of the elementary particles is the result of efforts of many scientists in fundamental physics over the last decades. These efforts in both theoretical and experimental aspects pushed the science borders forward. Nowadays, all elementary particles together with their interactions are described within a theory entitled "standard model of particles". This model is a *quantum field theory* which is the result of combining the quantum theory and the special relativity. The predictions of this theory have an extraordinary agreement with the experimental results and it is known as the most precise tested theory so far. Despite of its great success, the model has opened many questions in front of us. To answer these questions extending the model seems necessary.

In this chapter, the main concepts which the theory is based on are reviewed. In 1.1, the particle content of the Standard Model and the interactions are briefly explained. The emphasis is on the spontaneous breaking of the electroweak symmetry and the Higgs mechanism as one of the main motivations for the LHC construction. At the end of the section, the defects of this model are listed.

In the next section (1.2), Super Symmetry (SUSY) as one of the most favorable extensions to the Standard Model is introduced.

1.1 The Standard Model and Gauge symmetries

One of the most fundamental concepts in particle physics is that the interactions among elementary particles are described by symmetry principles. The invariance of the Lagrangian under certain symmetry transformations leads to a set of conservation laws and conserved quantum numbers (Nother theorem [1]).
The Standard Model of particle physics is a very successful description of fundamental interactions in terms of the gauge symmetries. The model unifies the electromagnetic and the weak forces using a $SU(2)_L \times U(1)_Y$ gauge interaction. While the $SU(2)_L$ symmetry describes the weak interaction between a lepton and a neutrino from the same generation, the $U(1)_Y$ symmetry is analogous to the electromagnetic U(1) interaction, happening due to the hypercharge Y of particles. The notation L emphasizes that only left-handed particles can participate in the weak interaction.
The electroweak symmetry is however broken by the so-called Higgs Mechanism [2, 3, 4] where the $SU(2)_L \times U(1)_Y$ symmetry changes to the electromagnetic U(1) symmetry

of the actual world,

$$\mathrm{SU(2)_L \times U(1)_Y \to U(1).}$$

The Higgs mechanism is also considered as the responsible for particles to become massive. Besides the broken electroweak symmetry, the Standard Model contains a non-broken color symmetry, $\mathrm{SU(3)_c}$, among quarks and gluons. Such interaction leads to the formation of nuclear particles, hadrons. The gravitational force which is by far weaker than the other fundamental interactions, is not included in this model. Quantum models for gravity suggest a spin-2 gauge boson, the graviton, to mediate the gravitational force. 1 summarizes the fundamental particles in the Standard Model. Three generations exist for the lepton and the quark sector where generations are defined according to their mass range. The heavier leptons are μ and τ together with their massless neutrinos. (c, s) and (t, b) are the members of the second and the third generations of quarks respectively. The weak force in both the quark and lepton sector is mediated by three massive gauge bosons ($\mathrm{W^{\pm}}$ and Z, $\mathrm{m} \sim 100\,\mathrm{GeV}$) while the electromagnetic (photon) and strong (gluon) gauge bosons are massless. The Higgs boson whose coupling with others introduces the mass term in the symmetry broken Lagrangian, is still not found.

Name	Symbol	$(SU(3), SU(2), U(1))$
Quarks		
Left-handed doublet	$(u, d)_L \equiv Q_L$	$(3, 2, \frac{1}{6})_L$
Right-handed up-type singlet	u_R	$(3, 1, \frac{2}{3})_R$
Right-handed down-type singlet	d_R	$(3, 1, -\frac{1}{3})_R$
Leptons		
Left-handed leptons	$(\nu_e, e^-)_L \equiv L_L$	$(1, 2, -\frac{1}{2})_L$
Right-handed charged leptons	e_R^-	$(1, 1, -1)_R$
Gauge bosons		
Gluons (strong force)	g	$(8, 1, 0)$
W and Z bosons (weak force)	$W^{\pm}, Z$	$(1, 3, 0)$
Photon (electromagnetic force)	γ	$(1, 1, 0)$
Scalar		
Higgs boson	H	$(1, 2, \frac{1}{2})$

Table 1: Particle content of the Standard Model. Of the three generations of quarks and leptons only the first is shown. Note that left-handed fermions are doublets under $SU(2)$ while right-handed fermions are singlets. $SU(2)$ and $SU(3)$ representations are labeled by their dimension while $U(1)$ representations are labeled by the eigenvalue of the $U(1)$ hypercharge generator Y.

1.1.1 Spontaneous Symmetry Breaking and the Higgs mechanism

Considering a left-handed lepton and its neutrino as a weak doublet,

$$L \equiv \begin{pmatrix} \nu_l \\ l \end{pmatrix}_L ,$$

we can write the conserved weak charged currents, $J_\mu^\pm$, as

$$J_\mu^\pm = \bar{L}\gamma_\mu \tau^\pm L,$$

where

$$\tau^\pm = \frac{\tau^1 \pm i\tau^2}{2}$$

and τ^i's are the known Pauli matrices. The third Pauli matrix, τ^3, induces the neutral current. These currents are further coupled to the gauge fields, A^μ, and form the Lagrangian of the weak interaction. Such Lagrangian does not include the electromagnetic interaction although the weak neutral current may seem to be a candidate for that. The $\mathrm{U(1)_Y}$ symmetry is introduced to the Lagrangian because unlike the weak neutral component, the electromagnetic force dose not discriminate between the left- and the right-handed particles and it is blind to neutrinos. This force is mediated by a gauge field B^μ and the hypercharge Y is related to the electromagnetic charge Q as

$$\frac{Y}{2} = Q - Q_w$$

where Q_w is the weak charge operator.
To incorporate the symmetry breaking in the model, the complex scalar fields of ϕ^0 (neutral) and ϕ^+ (positively charged) are introduced,

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix},$$

where they are described by a potential

$$V(\phi^\dagger\phi) = m^2(\phi^\dagger\phi) + \lambda(\phi^\dagger\phi)^2. \tag{1}$$

The interaction between these new fields and the fermions is described by a Yukawa

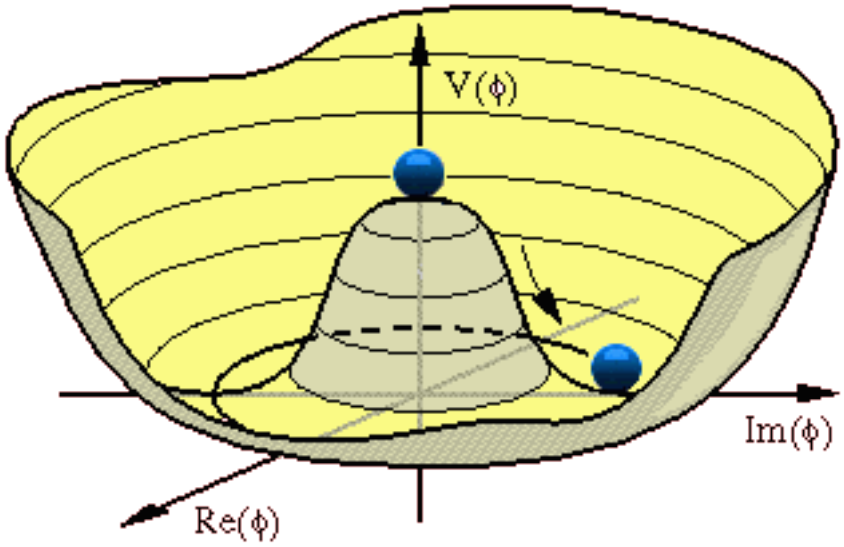


Figure 1: The Higgs potential.

interaction. For the gauge bosons, the coupling arises from the covariant derivative

$$D_\mu = \partial_\mu - \frac{iq_W g}{2}\vec{\tau}.\vec{A_\mu} - \frac{iYg'}{2}B_\mu. \tag{2}$$

The covariant derivative conserves the local electroweak gauge symmetry.

For a positive λ and a negative m^2, the potential in Equation (1) takes the shape as in Figure 1 on page 4 in the $\phi^0 - \phi^+$ space, hence a degenerate state for the minimum potential. The degeneracy or symmetry is broken if the potential is minimized by e.g. the scalar doublet of $\begin{pmatrix} 0 \\ v/\sqrt{2} \end{pmatrix}$ where

$$\phi^\dagger\phi = |\phi|^2 = v^2/2, \quad \text{with} \quad v = \sqrt{-m^2/\lambda}.$$

It can be shown that

$$Q_W\,\phi_0 = -\tfrac{1}{2}\phi_0, \quad \text{and} \quad Y\phi_0 = \phi_0,$$

but $Q\phi_0 = 0$. Since the vacuum state is not physical, it has to be annihilated by physical operators. This means that the ϕ_0 vacuum state preserves the electromagnetic symmetry, $U(1)$, while it breaks the $SU(2)_L$ and $U(1)_Y$ symmetries. Excitations of this

vacuum state can be parameterized as

$$\phi_0 = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = e^{i\vec{\tau}.\vec{\xi}/2v} \begin{pmatrix} 0 \\ (v+H)/\sqrt{2} \end{pmatrix},$$

where the phase term is further absorbed by a unitary transformation. The field $H(x)$ is the so-called Higgs field. Using this new representation in a trivial but lengthy calculation, mass terms for fermions together with the three massive gauge bosons ($W^{\pm}$,Z) and the massless photon will appear in the symmetry broken Lagrangian.

1.1.2 Shortcomings of the Standard Model

The Standard Model of particle physics is successful in explaining almost all aspects of the known particles and interactions. Due to a number of questions, the model is however thought as an effective theory of a more complete symmetry. In addition to the Higgs boson which is still not found, the fact that the gravitational force is not described by the model is one of its shortcomings. The theoretical desire of unifying all known forces in a simple symmetry is not accessible in the Standard Model even for the electroweak and the strong interactions. Another imperfection of the theory is the extremely large quantum correction to the Higgs mass known as the hierarchy problem. This leads either to a very large Higgs boson mass which prevents the theory to be perturbative or to fine tuning in calculations which is not favorable.
There are also evidences for the existence of massive neutrinos [5], not explained by the theory. Other experimental findings are dark energy[6] and dark matter[7, 8] which are responsible for a great amount of the matter/energy in the universe but not described by the model.
Towards an answer for such questions, different theories, including extensions to the Standard Model, are proposed. Among these models which need to be tested in high energy colliders such as LHC [9], the Super Symmetric Standard Model [10] provides a good explanation for the hierarchy problem. It gives a prospect for the unification of the strong and the electroweak forces with a *Grand Unified* theory at high energy ($\sim 10^{16} GeV$) while proposing a candidate for the invisible dark matter.

Hierarchy problem Considering the Planck mass scale

$$m_P = \sqrt{\frac{hc}{G}} \approx 10^{19}\,\text{GeV}$$

at which the quantum gravitational effects become important, the Standard Model cannot naturally explain why it is so much higher than the relatively low mass scale of the electroweak symmetry breaking ($\sim 10^2\,\text{GeV}$).
Also, for the theory to be perturbative, the Higgs boson mass is supposed to be below 1 TeV. However, the calculation of this Higgs boson mass is extremely sensitive to higher order quadratic and logarithmic divergences which occur as a result of radiative corrections. The logarithmic divergences are renormalized while the quadratic divergences cannot be regularized. The quantum contributions from the fermion loops into the Higgs mass take the form of

$$\Delta m_H^2|_F = \frac{|\lambda_F|^2}{16\pi^2}(-2\Lambda_{UV}^2 + 6m_F^2\, ln(\Lambda_{UV}/m_F) + ...), \tag{3}$$

where m_F is the fermion mass, λ_F is the coupling strength between the Higgs boson and the fermion, and Λ_{UV} is the Ultra-Violet momentum cut-off at which manifestations of new physics are expected. If there is no new physics between the electroweak and the Planck sale, Λ_{UV} has to be replaced with m_P which means a correction up to 30 orders larger than the expected Higgs mass, $10^2\,\text{GeV}$. The worst happens for the loop of top quark in which $\lambda_t \approx 1$.
To protect the Higgs mass and keep it positive, a finely tuned cancellation has to occur between the *bare* mass of the Higgs boson and its quantum corrections. It should be noted that the negative Higgs mass would not lead to the electroweak symmetry breaking. The same correction from the bosonic loops is

$$\Delta m_H^2|_B = \frac{|\lambda_B|}{16\pi^2}(2\Lambda_{UV}^2 + 2m_B^2\, ln(\Lambda_{UV}/m_B) + ...). \tag{4}$$

It can be seen that the sign of the quadratic divergence is positive hence giving the idea of a possible cancellation between the fermion and the boson corrections with $\lambda_B = \lambda_F^2$. Such cancellation implies a softly broken symmetry which is modeled in the *supersymmetry* framework. Alternative models like little Higgs[11] are also proposed to resolve the hierarchy problem. The focus in this thesis is however on the supersymmetry.

Unification of different forces The strength of the interaction in field theory is governed by the "coupling constant", α, which is not really a constant, but changes with the energy scale of the interactions, Q^2. The energy scale of the interaction is expressed with respect to a reference, Q^2_{ref}. It is shown in the framework of the renormalization group that the inverse of the coupling constant evolves linearly as a function of $\mathrm{t} = \log\left(\frac{Q^2}{Q^2_{ref}}\right)$, i.e.

$$\frac{d}{dt}(1/\alpha_i(t)) = -b_i/4\pi \qquad (i = 1, 2, 3) \tag{5}$$

The index i in Equation 5 stands for the three different interactions present in the Standard Model. It is desired to have $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_{GUT}$ at some energy scale (GUT energy) to unify different interactions.

The electromagnetic coupling constant, α_1, increases with the energy scale, given

$$\begin{aligned} b_1 = \ & \tfrac{4}{3}n_g + \tfrac{1}{10}n_h \\ & n_g = 3 \qquad \text{(number of generations)} \\ & n_h = 1 \qquad \text{(number of Higgs doublets)} \end{aligned}$$

The non-Abelian nature of the weak and strong interactions introduces an opposite behavior for their coupling constants, α_2 and α_3 respectively. They fall at higher energy scales since

$$\begin{aligned} b_2 &= -\tfrac{22}{3} + \tfrac{4}{3}n_g + \tfrac{1}{6}n_h \\ b_3 &= -\tfrac{11}{3} + \tfrac{4}{3}n_g \end{aligned}$$

Regarding these different behaviors, each of the strong and the weak couplings can be equal to the electromagnetic one at some energy scale. However, as illustrated in Figure 2 on page 8, there is no energy scale in the Standard Model framework at which all three couplings become equal. The existence of such unification is statistically excluded at a level of 12σ.

Thanks to the supersymmetry, the slope of the couplings evolution changes in such a way that the unification between three forces becomes possible.

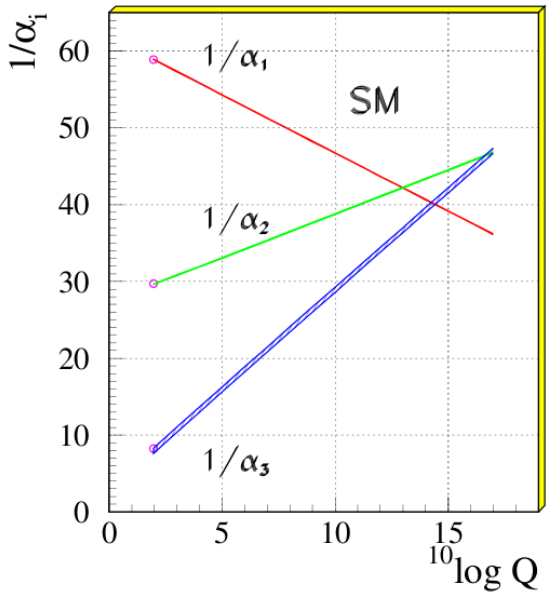


Figure 2: Running of the inverse of the three coupling constants in the Standard Model. It is seen that they do not cross at the same point.

1.2 Supersymmetry

Supersymmetry (SUSY) is a symmetry between fermions and bosons. It introduces a new operator $\mathcal{Q}$ to generate such transformations :

$$\mathcal{Q}|Boson\rangle = |Fermion\rangle$$

$$\mathcal{Q}|Fermion\rangle = |Boson'\rangle$$

In the simplest case which was developed by Wess and Zumino [12], we can consider the Lagrangian of a massless left-handed fermion (ψ) and a massless complex boson (Φ) :

$$\mathcal{L} = \partial^\mu \phi^* \partial_\mu \phi + i\psi^\dagger \overline{\sigma}^\mu \partial_\mu \psi \tag{6}$$

Then we can try to find the properties of the $\mathcal{Q}$ operator, so that the lagrangian (or the action which is $S = \int \mathcal{L} d^4x$) under its transformation remains invariant.

A clear manifestation of the $\mathcal{Q}$ operator is its fermionic properties. It appears in the

properties of the parameter of the transformation :

$$\delta_\epsilon \Phi = \epsilon\psi \qquad\qquad \delta_\epsilon \Phi^* = \epsilon^\dagger \psi^\dagger \tag{7}$$

ϵ which parametrizes the supersymmetry is a left-handed Weyl spinor and carries spin$-\frac{1}{2}$. The simplest possible transformation of the bosonic field preserving the invariance of Wess-Zumino action (6) is [10]:

$$\delta_\epsilon \psi_\alpha \;=\; i(\sigma^\mu \epsilon^\dagger)_\alpha \partial_\mu \phi, \qquad\qquad \delta_\epsilon \psi^\dagger_{\dot\alpha} \;=\; -i(\epsilon\sigma^\mu)_{\dot\alpha} \partial_\mu \phi^* \tag{8}$$

Now one can investigate the properties of the $\mathcal{Q}$ operator . The commutator of two supersymmetry transformations can be found by mixing the fermionic (7) and bosonic (8) transformations :

$$(\delta_{\epsilon_2}\delta_{\epsilon_1} - \delta_{\epsilon_1}\delta_{\epsilon_2})\phi \;\equiv\; \delta_{\epsilon_2}(\delta_{\epsilon_1}\phi) - \delta_{\epsilon_1}(\delta_{\epsilon_2}\phi) \;=\; i(-\epsilon_1\sigma^\mu\epsilon_2^\dagger + \epsilon_2\sigma^\mu\epsilon_1^\dagger)\partial_\mu\phi \tag{9}$$

The same result can be derived by applying two consecutive transformations on the bosonic field ψ, enforcing the equation of motion. The result becomes more clear after translating (9) in terms of the generators of the corresponding transformations :

$$\{\mathcal{Q}_\alpha, \mathcal{Q}^\dagger_\beta\} = 2\sigma^\mu_{\alpha\beta} P_\mu$$

This equation defines the supersymmetry algebra, which unlike the usual algebras are defined by the anticommutators of their generators.

A remarkable feature of the supersymmetry operators is that they connect the spin which is an internal symmetry of particles to the energy-momentum which is an external property.

This study can be extended to more generic results. The single-particle states of a supersymmetric theory fall naturally into irreducible representations of the supersymmetry algebra. These representations are called *supermultiplets*. Each supermultiplet contains both fermionic and bosonic states which are known as *superpartners* of each other. It can be easily shown that the mass and the fermionic and bosonic degrees of freedom should be equal for superpartners. In addition, the supersymmetry generators commute with the generators of gauge transformations. Therefore superpartners must also have the same charge, isospin and color.

One might wonder whether any of the known particles in the Standard Model are linked together in supermultiplets. The answer is no, because the known bosons have different internal quantum numbers than the known fermions. For example, no boson is known to match with the electric charge and color representation of the quarks.

Thus, to respect the supersymmetry, the only possibility seems to be to introduce new supersymmetric partners (spartners) for all SM particles. Then try to break the supersymmetry to describe the current situation of the particles.

The minimal supersymmetric extension of standard model is known as MSSM and is described and used in this thesis.

1.2.1 How SUSY solves the problems of Standard Model

In this section we will show why supersymmetry is the most attractive extension of the Standard Model. The defects of the Standard Model were outlined in Section 1.1.2. SUSY has its own solution for some of these problems.

Hierarchy The solution that SUSY provides for the hierarchy problem is one of its best motivations. The quadratic divergence due to loop corrections of the Higgs mass, which was discussed in Section 1.1.2, is naturally canceled in SUSY. Because for each boson which couples to Higgs boson, there is a fermion with the same mass. Since bosonic and fermionic loops have opposite, the residual one-loop correction to the Higgs $(\text{mass})^2$ is of the form:

$$\Delta m_H^2 = \frac{|\lambda_f|^2}{8\pi^2}(\frac{3}{2}m_f^2 - m_B^2)ln(\Lambda_{\text{UV}}/m_F) + \dots$$

One can show that the cancellation will also occur in higher order corrections.

Unification of gauge couplings The next success of supersymmetry is related to unification theory. It was shown that the coupling constants of electromagnetic, weak and strong interactions do not unify at a single point in the Standard Model. Adding new particles to the theory changes the slope of the evolution of the gauge couplings in (5).

It should be noted that the unification of the three curves at a single point is not at all trivial, since all three curves are affected simultaneously and the slopes are strongly

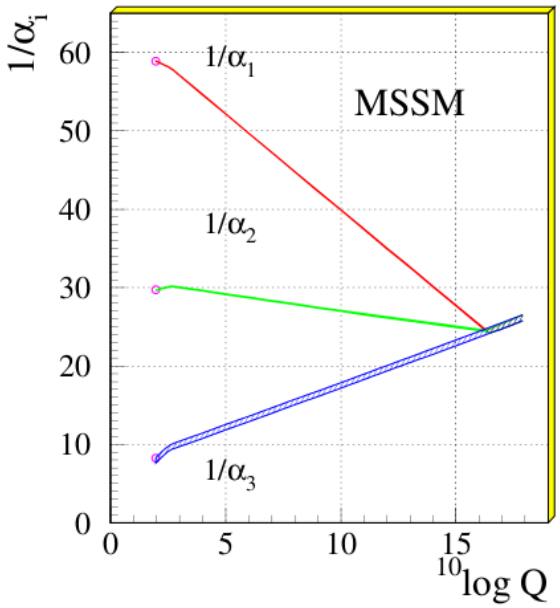


Figure 3: Running of the inverse of the three coupling constants in the MSSM. The thickness of the lines reflects the experimental uncertainty in the coupling constant determination.

correlated. For example adding new Higgs doublet or new generations of quarks never results in the unification of coupling constants.

The particle content of SUSY, which is twice the Standard Model particle content, changes the b_i coefficients to $(33/5, 1, -3)$. With these numbers a perfect unification of the gauge couplings is obtained if the sparticle masses are of the order of 1 TeV. This is shown in Figure 3 on page 11. The unified coupling constant α_{GUT} is about $1/26$ which is well inside the prohibitive regime and the unified theory is guarantied to be calculable .

Dark Matter The only non-relativistic, neutral, nonbaryonic particle in the Standard Model is Neutrino. It is too light to explain the observation of the existence of dark matter in the universe[13].

If SUSY preserves the R-parity, which is needed to explain the very long life time of protons ($\tau_{proton} > 10^{33} years$), the lightest supersymmetric particle (LSP) is always stable. So it can be a good candidate for the cold dark matter. Such a candidate is limited by several constraints as it should not be detectable. It must be electrically and color neutral so that they do not bind with normal matter and create detectable

isotopes.

1.3 The MSSM

In this section, we shortly review the particle contents of the minimal supersymetric extension of the Standard Model. Its Lagrangian together with the interaction between the particles will be described. Different scenarios to break the supersymmetry will be explained.

1.3.1 Supersymmetric Lagrangians

We want to extend the Standard Model with supersymmetry. The simplest way to acheive that is by introducing two kinds of supermultiplets:

- a *chiral supermultiplet*, which combines a Weyl fermion with a complex scalar.
- a *gauge supermultiplet*, which is the combination of a gauge boson with a Weyl fermion.

A complex scalar has two degrees of freedom, while an off-shell Weyl fermion has four. So we need to introduce an auxiliary field F, a complex scalar field that enters the Lagrangian as:

$$\mathcal{L}_F = F^* F \tag{10}$$

A similar adjustment is needed for the gauge supermultiplet, since an off-shell gauge boson has only three degrees of freedom compared to four degrees of freedom of an off-shell Weyl fermion. Introducing a real bosonic auxiliary field D^a solves this problem. The corresponding term in the Lagrangian is:

$$\mathcal{L}_D = \frac{1}{2} D^a D^a \tag{11}$$

The auxiliary fields are not physical, because they do not have kinetic terms. They are fixed by their equations of motion, which follow from the Lagrangians of the chiral and gauge supermultiplets.

Now we can use these building blocks to find the most general supersymmetric Lagrangian. We start with the chiral supermultiplets and find all of their possible supersymmetric interactions. We will show that these are specified by a superpotential. Then requiring gauge invarince, we can add the gauge supermultiplets and their interactions.

The Superpotential. It can be shown that the possible interaction terms for the chiral supermultiplet which preserve supersymmetry and keep the Lagrangian renormalizable, can be obtained using a superpotential W. W is an analytic function in the scalar fields ϕ_i and can be parametrized as :

$$W = L^i\phi_i + \frac{1}{2}M^{ij}\phi_i\phi_j + \frac{1}{6}y^{ijk}\phi_i\phi_j\phi_k \tag{12}$$

This leads to an interaction Lagrangian of the form [10]:

$$\mathcal{L}_{int} = \left(-\frac{1}{2}W^{ij}\psi_i\psi_j + W^iF_i\right) + c.c. \tag{13}$$

where W^{ij} and W^i are polynomials in the scalar fields that are related to each other as:

$$W^{ij} = \frac{\delta^2 W}{\delta\phi_i\delta\phi_j} \qquad W^i = \frac{\delta W}{\delta\phi_i}$$

The term with M^{ij} contains a fermion mass term, which is allowed in a general theory. Inserting equation 12 in the first term of the Lagrangian 13, a Yukawa interaction appears in the terms containing y^{ijk}. The linear term does not occur in the Standard Model.

Gauge Invariance. We can require the Lagrangian for the chiral supermultiplet to be gauge invariant. In the Standard Model, gauge invariance results in gauge bosons. In a supersymmetric theory we cannot introduce new particles without introducing their supersymmetric partners. So to make a gauge invariant Lagrangian, gauge supermultiplets that contain not only gauge bosons, but also gauginos are added. The D terms also should be considered in the Lagrangian.

Writing all of the possible interactions, the F and D terms are fixed by their equations of motion. It turns out that [10]:

$$F_i = -W_i^* \qquad F^{*i} = -W^i \qquad D^a = -g\phi^*T^a\phi \tag{14}$$

where T^a is the generator of the gauge group and g the gauge coupling.
The important thing is that the Lagrangian is completely determined by the superpotential and the gauge groups. The superpotential specifies the chiral supermultiplets and their interactions. By requiring gauge invariance we can derive the gauge super-

Table 2: Chiral supermultiplets in the Minimal Supersymmetric Standard Model. The spin-0 fields are complex scalars, and the spin-1/2 fields are left-handed two-component Weyl fermions.

Names		spin 0	spin 1/2	$SU(3)_C$, $SU(2)_L$, $U(1)_Y$
squarks, quarks (×3 families)	Q	$(\widetilde{u}_L \;\; \widetilde{d}_L)$	$(u_L \;\; d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	$\bar{u}$	$\widetilde{u}_R^*$	$u_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	$\bar{d}$	$\widetilde{d}_R^*$	$d_R^\dagger$	$(\overline{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons (×3 families)	L	$(\widetilde{\nu} \;\; \widetilde{e}_L)$	$(\nu \;\; e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	$\bar{e}$	$\widetilde{e}_R^*$	$e_R^\dagger$	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	H_u	$(H_u^+ \;\; H_u^0)$	$(\widetilde{H}_u^+ \;\; \widetilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	H_d	$(H_d^0 \;\; H_d^-)$	$(\widetilde{H}_d^0 \;\; \widetilde{H}_d^-)$	$(1, \mathbf{2}, -\frac{1}{2})$

multiplets and all other interactions.

1.3.2 The Minimal Supersymmetric Standard Model

Although the basic form of a supersymmetric Lagrangian is derived, this still needs many free parameters to construct an actual theory. One can introduce any number of chiral supermultiplets and gauge symmetries or even have several supersymmetry generators. We want to restrict ourselves to the simplest extension of the Standard Model, with one supersymmetric generator and a minimal particle content.

Particle Content In the simplest case, we can assign a scalar superpartner to each fermion. In Table 2 on page 14, all of the superchiral multiplets are listed.

The Higgs boson fields as the sole scalar fields in the Standard Model, make a chiral supermultiplet with spin ½ particles named Higgsinos.

Two Higgs Doublets. In supersymmetry we need two Higgs doublets. We know that the charge conjugate of the Standard Model Higgs doublet is used to give mass to the up-type fermions. However, the charge conjugate of a scalar field is not an analytic function in that field, so it cannot appear in the superpotential. Therefore a second Higgs doublet with opposite hypercharge is introduced to give mass to the down-type fermions after spontaneous symmetry breaking. Explicitly, the Higgs doublets are:

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \qquad H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix}$$

Table 3: Gauge supermultiplets in the Minimal Supersymmetric Standard Model.

Names	spin 1/2	spin 1	$SU(3)_C, SU(2)_L, U(1)_Y$
gluino, gluon	$\widetilde{g}$	g	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\widetilde{W}^{\pm}$ $\widetilde{W}^0$	$W^{\pm}$ W^0	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	$\widetilde{B}^0$	B^0	$(\mathbf{1}, \mathbf{1}, 0)$

Only the neutral components H_i^0 can get a vev, for otherwise the photon would couple to the charged vev and acquire a mass. So it would seem that the H_2 doublet gives mass to the down-type fermions.

The gauge sector of the Standard Model is also extended. As usual, a tilde is used to denote the supersymmetric partner of a Standard Model state, so the symbols for the gluon and gluino are g and $\widetilde{g}$ respectively. The electroweak gauge symmetry $SU(2)_L \times U(1)_Y$ is associated with spin-1 gauge bosons W^+, W^0, W^- and B^0, with spin-1/2 superpartners $\widetilde{W}^+, \widetilde{W}^0, \widetilde{W}^-$ and $\widetilde{B}^0$, called *winos* and *bino*. After electroweak symmetry breaking, the W^0, B^0 gauge eigenstates mix to give mass eigenstates Z^0 and γ. The corresponding gaugino mixtures of $\widetilde{W}^0$ and $\widetilde{B}^0$ are called zino ($\widetilde{Z}^0$) and photino ($\widetilde{\gamma}$); if supersymmetry were unbroken, they would be mass eigenstates with masses m_Z and 0. Table 3 on page 15 summarizes the gauge supermultiplets of a minimal supersymmetric extension of the Standard Model [14].

Super potential The Minimal Supersymmetric Standard Model (MSSM) is specified by the superpotential:

$$W = \tilde{u}\mathbf{y}_u\tilde{Q}H_2 - \tilde{\tilde{d}}\mathbf{y}_d\tilde{Q}H_1 - \tilde{e}\mathbf{y}_e\tilde{L}H_1 + \mu H_2 H_1 \tag{15}$$

The Yukawa couplings $\mathbf{y}_i$ are 3×3 matrices in the family space of generations and all gauge and family indices have been suppressed.

The parameter μ is the only new parameter so far that is added to the Standard Model. It is needed since $\mu = 0$ is excluded by the LEP experiments [15]. μ is called the Higgs mixing parameter or the higgsino mass parameter. The μ-term is unique, because terms $H_u^* H_u$ or $H_d^* H_d$ are not allowed in the superpotential because of its analyticity [14].

1.3.3 Supersymmetry Breaking

Unbroken supersymmetry predicts that there are no mass differences within supermultiplets. Clearly we need to find a way to break supersymmetry. Several methods have been proposed, but none of them is able to derive a broken supersymmetric Lagrangian from first principles.

Breaking Mechanisms The only known way to break a symmetry while maintaining the characteristics of a theory is spontaneous symmetry breaking. So our only hope is to do something similar and change our theory in such a way that the vacuum is not invariant under the supersymmetry transformation $\mathcal{Q}$. There are two terms in the supersymmetry Lagrangian that can acquire a vev without breaking other symmetries: the F terms and the D terms. Using the D terms does not work and in addition to the supersymmetry breaks the gauge symmetries of the theory. To use the F terms, we need to introduce a "hidden-sector" that contains particles that couple very weakly to the particles we observe. These small interactions would be responsible for symmetry breaking through radiative corrections. There are many possible scenarios for this hidden sector and the way it couples to the visible sector.

If the hidden and the visible sectors interact with each other via gravitational interactions, the model is called SUGRA.

Soft Breaking Terms Although no simple mechanism exists to break the supersymmetry, we can add new terms to the Lagrangian to break it in the low energy and vanish at high energies. The possible terms which are called soft breaking terms are [10]:

- Gaugino masses for each gauge group: $-\frac{1}{2}(M_3\tilde{g}\tilde{g} + M_2\tilde{W}\tilde{W} + M_1\tilde{B}\tilde{B} + c.c.)$ with $M_i \in \mathbb{C}$
- Sfermion masses: $-\tilde{Q}^\dagger \mathbf{m}^2_{\tilde{Q}}\tilde{Q} - \tilde{L}^\dagger \mathbf{m}^2_{\tilde{L}}\tilde{L} - \tilde{\bar{u}}\mathbf{m}^2_{\tilde{\bar{u}}}\tilde{\bar{u}}^\dagger - \tilde{\bar{d}}\mathbf{m}^2_{\tilde{\bar{d}}}\tilde{\bar{d}}^\dagger - \tilde{\bar{e}}\mathbf{m}^2_{\tilde{\bar{e}}}\tilde{\bar{e}}^\dagger$ where the mass matrices are 3×3 hermitian matrices in family space
- Higgs masses and mixing: $-m^2_{H_2}H_2^\dagger H_2 - m^2_{H_1}H_1^\dagger H_1 - (bH_2H_1 + c.c.)$ with $m^2_{H_i} \in \mathbb{R}, b \in \mathbb{C}$
- Triple scalar couplings: $-(\tilde{\bar{u}}\mathbf{a}_u\tilde{Q}H_2 - \tilde{\bar{d}}\mathbf{a}_d\tilde{Q}H_1 - \tilde{\bar{e}}\mathbf{a}_e\tilde{L}H_1 + c.c.)$ where the scalar couplings are complex 3×3 matrices in family space.

These terms break supersymmetry, but they conserve R-parity. Also, in contrast to the equivalent mass terms for the Standard Model particles, the gaugino and sfermion mass terms respect chiral and gauge symmetry.

We introduced supersymmetry at the cost of a single new parameter (μ). Including the soft breaking terms leaves us with over a hundred free parameters. That means our theory has lost all predictive power. Fortunately, experimental limits put stringent bounds on many of them. The experimental constraints on breaking parameters are summarized in 4.

Table 4: Some experimental constraints on the form of soft supersymmetry breaking parameters. The last column mentions the most important sources of these constraints.

Breaking term	Constraint	For instance constrained by
$\mathbf{m}^2_{\tilde{e}}, \mathbf{m}^2_{\tilde{L}}, \mathbf{a}_e$	Small off-diagonal elements	$\mu^- \to e^- \gamma$
$\mathbf{m}^2_{\tilde{u}}, \mathbf{m}^2_{\tilde{Q}}, \mathbf{a}_u$	Small off-diagonal elements	$D^0 \leftrightarrow \bar{D}^0$ mixing
$\mathbf{m}^2_{\tilde{d}}, \mathbf{m}^2_{\tilde{Q}}, \mathbf{a}_d$	Small off-diagonal elements	$K^0 \leftrightarrow \bar{K}^0$ mixing
	Complex phases have to be small	CP-violation

These constraints can be summarized as the hypothesis that all mass matrices are proportional to the unit matrix, that the triple scalar couplings are proportional to the Yukawa matrices and that breaking parameters introduce no complex phases [14].

1.3.4 Spontaneous Symmetry Breaking

For spontaneous symmetry breaking the minimum of the Higgs potential should occur away from the origin. This requirement is not met in unbroken supersymmetry, but the soft breaking terms change the potential. Writing out the Higgs fields in their components, the Higgs potential with the breaking terms included is:

$$V = \left(|\mu|^2 + m^2_{H_2}\right)\left(|H_2^0|^2 + |H_2^+|^2\right) + \left(|\mu|^2 + m^2_{H_1}\right)\left(|H_1^0|^2 + |H_1^-|^2\right) + \left[b(H_2^+ H_1^- - H_2^0 H_1^0) + c.c.\right]$$
$$+ \frac{1}{8}\left(g^2 + g'^2\right)\left(|H_2^0|^2 + |H_2^+|^2 - |H_1^0|^2 - |H_1^-|^2\right)^2 + \frac{1}{2}g^2\left|H_2^+ H_1^{0*} + H_2^0 H_1^{-*}\right|^2$$

where the breaking parameters b and m_{H_i} can take complex values. The equation can be simplified by gauge fixing. First using $SU(2)$ rotations we can choose $H_2^+ = 0$. Since in a minimum we must have $\partial V / \partial H_2^+ = 0$, this also implies $H_1^- = 0$. Then we use the fact that the b term is the only term that depends on the phase of the Higgs fields.

Using a $U(1)$ phase rotation to make b real and positive, it follows that in a minimum of the potential $H_2^0 H_1^0$ must be real and positive as well. So: the Higgses have opposite phases. Since they have opposite hypercharge, we can choose both Higgses to be real and positive. Choosing this gauge the equation takes a simpler form:

$$V = (|\mu|^2 + m_{H_2}^2)|H_2^0|^2 + (|\mu|^2 + m_{H_1}^2)|H_1^0|^2 - [bH_2^0 H_1^0 + c.c.] + \frac{1}{8}(g^2 + g'^2)(|H_2^0|^2 - |H_1^0|^2)^2$$

The quartic terms ensure that the potential is bounded from below for almost all values of H_2^0 and H_1^0. However for $H_2^0 = H_1^0$ the quartic terms vanish and the potential will not be bounded from below unless:

$$2b < \left(|\mu|^2 + m_{H_2}^2\right) + \left(|\mu|^2 + m_{H_1}^2\right) \tag{16}$$

This implies that not both $|\mu|^2 + m_{H_1}^2$ and $|\mu|^2 + m_{H_2}^2$ can be negative, so we can only have a minimum outside the origin if the origin itself is a saddle point, which is the case if:

$$b^2 > (|\mu|^2 + m_{H_2}^2)(|\mu|^2 + m_{H_1}^2) \tag{17}$$

If equations 16 and 17 are met, supersymmetry breaking results in spontaneous symmetry breaking. The ratio of the vevs of the neutral Higgs components is usually written as:

$$\tan\beta = \frac{\langle H_2 \rangle}{\langle H_1 \rangle} \tag{18}$$

Like in the Standard Model, the Higgs vevs are related to the mass of the W and Z bosons m_W and m_Z. The minimum of the potential satisfies $\partial V/\partial H_1 = \partial V/\partial H_2 = 0$, which gives us two equations we can use to express b and $|\mu|^2$ in terms of $\tan\beta$. At tree level, the result is:

$$b = \frac{m_{H_1}^2 - m_{H_2}^2 + m_Z^2 \cos(2\beta)}{\tan\beta - \cot\beta} \tag{19}$$

$$|\mu|^2 = \frac{1}{2}\left(b(\cot\beta + \tan\beta) - m_{H_1}^2 - m_{H_2}^2\right) \tag{20}$$

A complex phase of μ would introduce large CP-violating effects that are not observed in nature, so we demand μ to be real. That means we can eliminate b and μ in favour of $\tan\beta$ and the sign of μ. For phenomenological purposes, these variables are more convenient, since particle masses are closely related to $\tan\beta$.

In the Standard Model, three of the four degrees of freedom of the Higgs doublet were 'eaten' by the gauge bosons and one physical Higgs boson was left. In the MSSM there are two complex Higgs doublets, so a total of eight degrees of freedom. The gauge bosons still absorb three of them, so we now end up with five Higgs bosons: h^0, H^0, H^+, H^- and A^0. By convention h^0 is lighter than H^0 and A^0 is the only CP-odd state.

Due to the mixing terms, the mass eigenstates are considerably different from the gauge eigenstates. Diagonalizing the mass matrix gives the following masses at tree level [10]:

$$m_{A_0}^2 = \frac{2b}{\sin(2\beta)} \tag{21}$$

$$m_{h^0,H^0}^2 = \frac{1}{2}\left(m_{A^0}^2 + m_Z^2 \mp \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)}\right) \tag{22}$$

$$m_{H^\pm}^2 = m_{A^0}^2 + m_W^2 \tag{23}$$

This result implies a low h^0 mass. In fact $m_{h^0} \leq m_Z|\cos(2\beta)|$ [10], which is excluded experimentally. Due to large radiative corrections from especially tops and stops, the actual h^0 mass is larger, but it cannot exceed 135 GeV [16], which means it has to be found at the LHC if supersymmetry exists. The other Higgs bosons are much heavier and nearly degenerate in mass.

1.3.5 Minimal gravity mediated super symmetry breaking (mSUGRA)

mSUGRA is one of the most widely used models in supersymmetry. It assumes that breaking occurs through a coupling to gravity in its simplest form. Since gravity is flavour-blind, this justifies the assumption that the breaking mass matrices are proportional to the unit matrix. It also assumes universality of the scalar masses and of the gaugino masses at a high energy scale. This leaves just five free parameters: $M_0, M_{1/2}, A_0, \tan\beta$ and the sign of μ, where:

$$M_{1/2} = M_1 = M_2 = M_3 \tag{24}$$

$$M_0^2 = \mathbf{m}_{\tilde{Q}}^2 = \mathbf{m}_{\tilde{L}}^2 = \mathbf{m}_{\tilde{\bar{u}}}^2 = \mathbf{m}_{\tilde{\bar{d}}}^2 = \mathbf{m}_{\tilde{\bar{e}}}^2 \tag{25}$$

$$M_0^2 = m_{H_1}^2 = m_{H_2}^2 \tag{26}$$

$$\mathbf{a}_u = A_0 \mathbf{y}_u \quad \mathbf{a}_d = A_0 \mathbf{y}_d \quad \mathbf{a}_e = A_0 \mathbf{y}_e \tag{27}$$

mSUGRA involves two assumptions that are weakly motivated. Even in the Standard Model we know of small nonzero parameters, such as the off-diagonal components and the CP-violating phase in the quark mixing matrix. So the case for strict soft supersymmetry breaking universality is not very strong. Yet even though it might not be exact, the constraints in table 4 show that soft supersymmetry breaking universality is a good approximation.

The second assumption is unification. The MSSM is consistent with the minimal GUT. In fact, the most important reason for assuming unification is a practical one. As long as no signs of supersymmetry have been found, the main purpose of models is to reveal general phenomenological implications. It is conceivable that dropping the requirement of unification yields a completely different phenomenology, but it is not very likely. Assuming unification reduces the complexity of the problem, but still allows a broad range of phenomenological scenarios.

In short, the most important property of mSUGRA is its predictive power. A theory with too many parameters is simply impractical, so we need the additional assumptions to make supersymmetry a workable theory.

The phenomenological and experimental studies on mSUGRA are usually reported in 2 dimensional plane of $m_0 - m_{1/2}$ where three other parameters are fixed.

The contour plots of squark and gluino masses in different $m_0 - m_{1/2}$ planes are shown in Figure 4 on page 21. The same plot for the lightest chargino and slepton masses are also presented in Figure 5 on page 21.

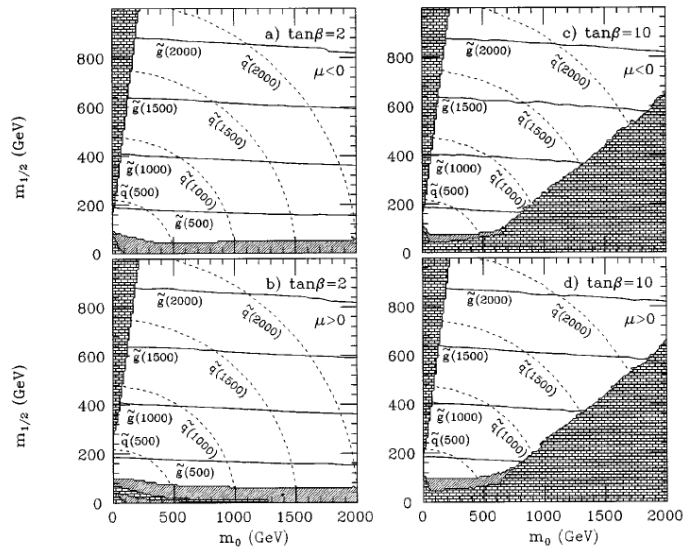


Figure 4: Contour plots of squark and gluino masses in the $m_0 - m_{1/2}$ plane of the minimal SUGRA model.

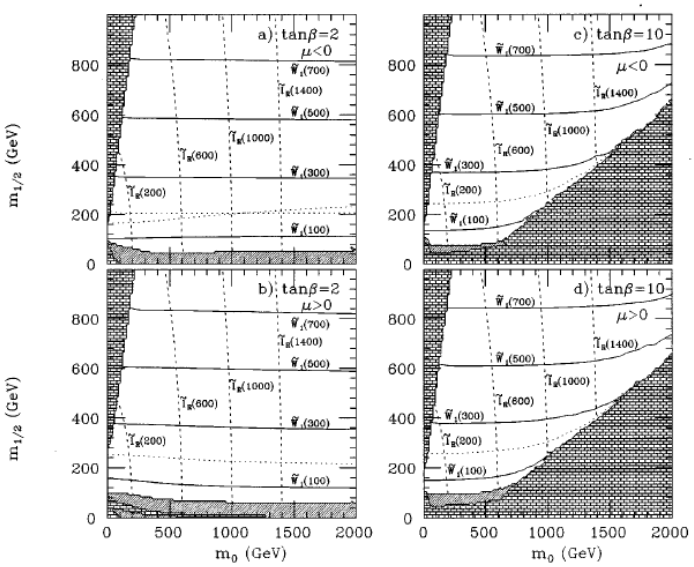


Figure 5: Contours of lightest chargino mass and contours of right slepton mass in the $m_0 - m_{1/2}$ plane of the minimal SUGRA model.

2 The Large Hadron Collider (LHC)

In a tunnel near the Franco-Swiss border, the Large Hadron Collider (LHC) has been built to help physicists to address the main important questions in the subatomic world. As discussed in 1.2, the Standard Model of particles can predict almost every subatomic phenomenon that has been observed thus far very well, although it has some key predictions for the origin of mass and the source of electroweak symmetry breaking which has not shown any signal yet. Moreover from a theoretical point of view it has some shortages that are described in details in Section 1.1.2.

This section reviews the history of colliders before LHC in brief then goes into details of LHC specifications and design goals. Then the 2008 accident which affected the schedule and physics programs is reviewed. In the last part the results of 2010 fills and the beam intensity and energy evolution during 2010 are discussed. The plan of the early future, till 2012 is also quoted there.

2.1 Colliders in general

Colliders are built to investigate the fundamental interactions and the finest constituents of matter. They are generally categorized by the type of particles they accelerate. Hadron colliders, like LHC and Tevatron, use proton or antiproton. The heaviness of these particles reduces the rate of losing energy and it makes the process of acceleration in a synchrotron easier. Protons and antiprotons are not elementary particles and are made up of quarks and gluons. Even if the two beams have the same energy, each constituent carries just a fraction of the beam energy. So in hadronic colliders the longitudinal momentum of the interaction is unknown and also the effective energy of the collision varies in each collision. In addition a strong interaction between two hadronic beams are more probable than other interactions. It makes the hadron colliders a good factory for QCD studies, but separating other interesting events from the sea of QCD events is not an easy task. In general, hadron colliders are not mainly built for a precise measurement but they can explore a wider energy region for new physics.

On the other hand lepton colliders, like LEP, are the best machines for very precise measurements. They usually collide electrons and/or positrons, although there are some proposals for Muon colliders too[17]. Hadron-Electron colliders like HERA in DESY also have been built to investigate the structure of protons in detail.

Among many other parameters that distinguish different colliders, energy and luminosity are the most important ones. The importance of the center of mass energy of the collision is clear, the more energy in a collision we can reach the more unknown spaces we can explore.

The interaction rate is proportional to the luminosity of the collider. There are two beams circulating in the accelerator, each beam consists of bunches of particles. In a head-on collision, we have

$$\mathcal{L} = \frac{fn^2}{4\pi\sigma_x\sigma_y} \tag{28}$$

In this formula, f is the frequency of the collisions. As the speed of bunches is very close to the speed of light f is closely related to the number of bunches per beam. n is number of protons per bunch, here we assumed that bunches in different beams have the same number of protons. σ_x and σ_y also represents the transverse size of the bunches at collision point. This formula needs a correction factor if there is a crossing angle between beams.

The beam size (the denominator in the luminosity formula) can be expressed in terms of other quantities : the transverse emittance,ϵ and the amplitude function at the interaction point, β^*. Both of these variables have units of length. The emittance reflects the process of bunch preparation back to the source of hadrons, The lower emittance of a particle beam the smaller spatial and momentum spread it has. On the other hand, the amplitude function (β) is determined by the accelerator magnet configuration and power, mostly the quadrapole magnet arrangement. In the detectors and before two beams collide, the beams are squeezed in order to produce more collisions. β^* is the length in which the beam size is reduced by a factor of 2 before reaching its minimum size, hence it is very correlated with the magnet power used for squeezing. A schematic view of the squeezing and the meaning of β^* can be seen in Figure 6 on page 25 [18].

The luminosity formula can be rewritten using these two variables as :

$$\mathcal{L} = \frac{fn^2}{4\epsilon\beta^*} \tag{29}$$

The instantaneous luminosity has the dimension of $[area]^{-1}[time]^{-1}$. The instantaneous luminosity though is not interesting by itself, its integral over time contains physical meaning.

$$L = \int \mathcal{L} dt$$

In high energy physics, *barn* (which is denoted by b) is usually used instead of $meter^2$ as the area dimension. Each barn is $10^{-28}m^2$ or $10^{-24}cm^2$. The instantaneous luminosity is sometimes expressed as $cm^{-2}s^{-1}$. The integrated luminosity which is a measure of the amount of data which is produced by a collider is in the units of $area^{-1}$ and for large colliders like LHC is usually expressed as $(fb)^{-1} = 10^{15}b^{-1}$ or $(pb)^{-1} = 10^{12}b^{-1}$.

A collision can lead to either *an elastic* or *inelastic* scattering. In an elastic scattering, the total kinetic energy of the particles doesn't change. When two beams pass through each other without any change in the structure of their particles, and the interactions between particles causes a change in the path of the particles, an elastic scattering has happened. If the momentum transfer is lower, the electromagnetic interaction is most important. For high momentum transfers, the QCD interactions are also involved.

But in the cases where two particles from two different beams get so close to each other that the internal structure of the particles can be seen or a new particle is created, an inelastic scattering has occurred. This second type of scattering is of interest to high energy physics.

The cross section of different kinds of scatterings can be calculated from the Lagrangian of a given theory or model. To obtain the expected number of a specific process in a data set with integrated luminosity of L, we just need to multiply the cross section of that process by the L:

$$< N_{expected} >= \sigma \times L \tag{30}$$

The cross section has units of area and the result is dimensionless as expected. Indeed the number of expected events are not fixed and has a Poisson distribution with the mean given by (30).

2.1.1 LEP

The large hadron collider has been built in the tunnel of LEP, the previous accelerator at CERN which used to collide electrons and positrons. The Large Electron Positron collider (LEP) is the most powerful accelerator of leptons ever built. As it was spe-

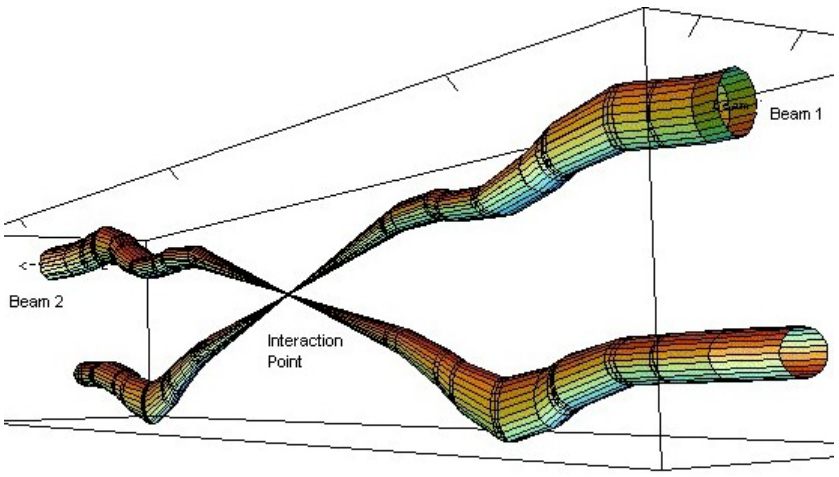


Figure 6: Schematic view of $\beta*$ and the squeezing of the beams before collision.

cially designed for the accurate measurement of Z mass, which had been discovered less than 10 years before LEP startup, it started with a center of mass energy of about 90GeV. The LEP tunnel excavation started in 1983 and the 27 km tunnel about 100 m underground was ready after less than 5 years in 1988. In mid 1989 LEP started to collide electrons and positrons and all the experiments installed started to take and record data. Four experiments were installed in LEP : ALEPH, DELPHI, OPAL and L3. Until 2000 and after 11 years of successful data taking, LEP energy was upgraded several times, first it was doubled to 189 GeV to have enough energy to produce WW events to measure W properties too. Then more superconducting cavities were added and the energy could be increased up to 209 GeV. LEP experiments could measure Z[19] and W[**ALEPH_W_Mass**, 20] boson properties very precisely. Another outstanding result from LEP is that only three types of low mass neutrinos exist and hence only three generations of lepton particles can exist in the world[21, 22]. In addition to the very precise standard model tests and measurements, new bounds on new physics and specially SUSY were found and reported[23]. In the last year of running, a small excess which was reported by all of the LEP experiments could postpone the deadline of LEP only by two months and in November 2000 LEP closed in favor of LHC construction.

2.1.2 Tevatron

Tevatron, the predecessor of LHC, is a hadron collider at the Fermi National Accelerator Laboratory in the United States of America. It is accelerating protons and antiprotons

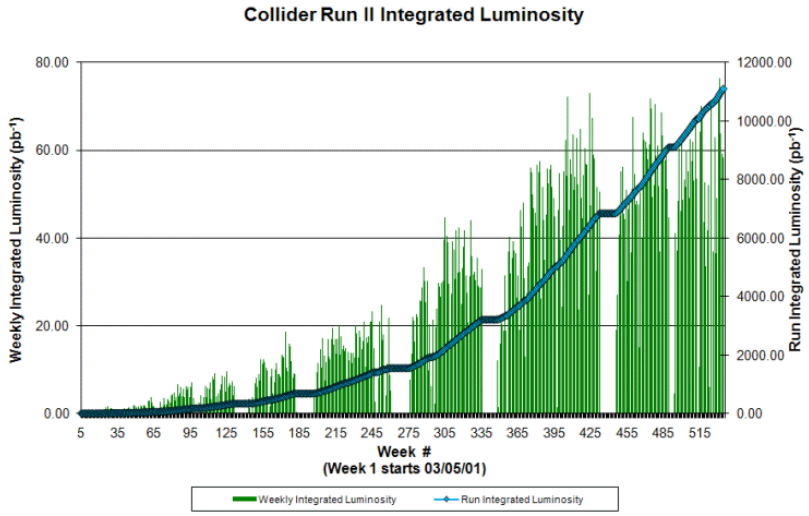


Figure 7: The weekly and total integrated luminosities for Run II of Tevatron. This plot is taken from the Tevatron Luminosity web site.

in a 6.28km ring. The synchrotron and the injectors were upgraded several times since 1983, when Tevatron started data taking, in order to improve and increase the energy of the beams to be able to scan new physics domains and also the beam intensity and collision luminosity. During the Run I collider program which began in 1992 and lasted for 4 years, the collision energy was 1.8TeV. After a five-year hiatus, Tevatron started Run II with 10% increase in energy : the center of mass energy was increased to 1.96TeV. Run II continues up to now and more than 11 fb^{-1} of data is recorded by the experiments, the luminosity plot of Run II can be seen in 7. The plan is to terminate the operation of Tevatron at the end of September 2011.

There are two main experiments to record the Tevatron collision data : D0 and CDF. Collider Detector at Fermilab (CDF) which is taking data since 1985 is the first detector installed at the Tevatron. During the first run of Tevatron, both experiments could record about 150 pb^{-1} of data and it was enough to discover the top quark [24, 25], the heaviest matter particle predicted by the standard model physics.

Recently, on April 2011, CDF experiment reported an excess on the invariant mass of the jets associated with a W boson in W+jj events [26]. This can possibly be the signature of a new particle, although D0 has not observed anything in that region yet [27].

2.2 LHC

2.2.1 Motivations for LHC

Very rare phenomena need colliders with ultra high luminosity to be discovered. LEP and Tevatron could not find any signature beyond the predictions of standard model of particles and Higgs particle is still hidden. Hence we need a machine with very high luminosity to discover the nature of beyond standard model.

The energy is also important. As it was already planned to use LEP tunnel for LHC, the dimensions of LHC were already fixed : it has to fit in a circular 27km tunnel. The energy loss due to synchrotron radiation is given by :

$$\Delta E = \frac{4\pi\alpha}{3R}\,\beta^3\gamma^4 \tag{31}$$

In the ultra-relativistic regime where our accelerators work β and γ are

$$\beta = \frac{v}{c} \approx 1,\ \gamma = \frac{E}{mc^2} \tag{32}$$

Putting (32) in (31) and simplifying the equations, we will have :

$$\Delta E \sim \frac{E^4}{Rm^4} \tag{33}$$

It means with a fixed radius, $R \approx 8.4Km$ in the case of LHC, the energy loss due to the synchrotron radiation increases proportional to E^4. i.e. if we want to increase the center of mass energy of the collisions by a factor of 10, ΔE would be 10000 times larger !To overcome this problem, there is another parameter to change : m, the mass of the accelerating particle. Increasing the mass of the particle can reduce the synchrotron radiation significantly.

An obvious choice was proton and/or antiproton which is still being used in Tevatron. A proton is 200 times heavier than electron. Hence with the same amount of energy loss, we can increase the energy by a factor of 200. This it is not what happened in reality due to lots of other technical restrictions.

A strong magnetic field is needed to bend the beams of protons. The relativistic relation between magnetic field and the energy of the accelerating particle in the LHC tunnel is given by :

$$E\,[TeV] = 0.84 \times B\,[Tesla] \tag{34}$$

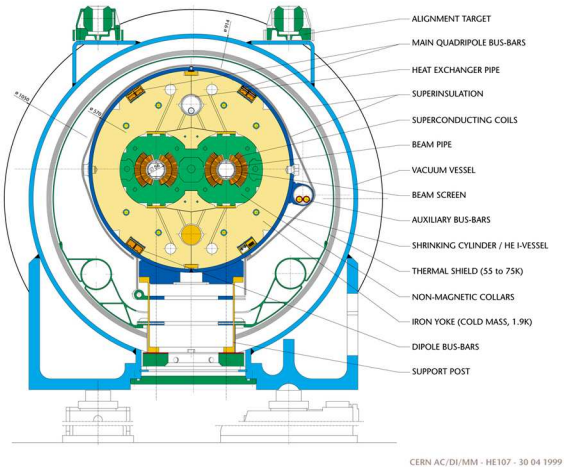


Figure 8: Schematic view of the LHC dipole magnet design

The new physics phenomena are expected to be around 1 TeV, so the new accelerator should scan up to $\approx 1TeV$. Electrons and generally leptons are elementary particles, on the contrary the protons are composite particles and each part of them carry only a fraction of the energy of the proton. So to reach 1TeV collisions in a proton collider, the energy of the protons should be in the range of multi TeV. Considering all technical restrictions at the time of designing LHC, dipole magnets with a 8.33 T nominal magnetic field were proposed and are installed in the LHC. According to (34), it is equivalent to a 7TeV beam i.e. the collisions with the center of mass energy of 14TeV.

To reach higher luminosities, more particles should be injected in the collider. Producing anti-proton beam is not as easy as proton, which is simply a beam of positively charged hydrogen ions. To be able to increase the number bunches and number of particles per bunch which directly affects the luminosity (see formula(28)), LHC designed to accelerate two proton beams instead of what Tevatron does. This decision has pros and cons : it makes higher luminosities possible by dealing with two proton beams in the accelerator. In the accelerators with two opposite charge beams, one magnet system is enough. For LHC two different pipes and dipole magnets are needed. The cross section of LHC dipole magnets can be seen in 8. Then two beams intersect at the experimental interaction points.

To prepare a magnetic field up to 8.33 Tesla, a current around 12kA is needed. NbTi superconductor is used to achieve this high electrical current. Super fluid helium at 1.9K is used to cool the NbTi alloy coil down to the super conductive state. The

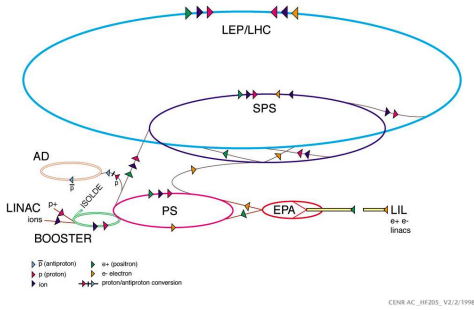


Figure 9: Schematic view of the CERN accelerator complex

energy of protons when injected in the LHC tunnel is 450GeV and during 20' during which the energy of the beams is increased up to 3500GeV, the magnetic field has to be ramped up from 0.53T to 8.33 Tesla [9].

Before injection into the LHC storage ring, the protons should pass a long way along older and simpler accelerators. The first step is LINAC2, a linear accelerator. Then the 50MeV protons from LINAC2 are injected into PSB (Proton Synchrotron Booster) and the booster accelerates them to 1.4GeV. In the PS the protons are accelerated up to 25GeV and in the last step before LHC, SPS (Super proton synchrotron) increases their energy by another factor of 18. SPS has had an important role in the history of particle physics when in 1982 it was used as proton-antiproton collider and the data provided by its beams to UA1 experiment led to the W and Z discovery [28, 29]. The 450GeV proton beams are finally transferred into LHC, clockwise and counterclockwise. The filling time is 4'20" and LHC will accelerate them up to the nominal energy [30, 9]. The chain of the CERN accelerator complex can be seen in 9.

Considering the 25 ns time interval between bunches in the LHC beam structure, the LHC ring has the capacity of 3564 bunches per beam. At different stages of injection, the "kickers" will leave gaps of missing bunches in the beam. Due to this reason and some other technicalities, LHC will be able to contain only 2808 bunches per beam at the same time. Each bunch consists of $\sim 10^{11}$ protons. It means that each beam in LHC has $\sim 3 \times 10^{14}$ protons which is $\sim 10^{-10}$ grams of matter and in LHC their energy would raise to 350MJ which is equivalent to the energy of 150 Kg of TNT.

2.2.2 2008 incident

After a long wait and some subsequent schedule changes, LHC decided to start in September 2008. But on the 19th and during powering tests of the main dipole circuit in Sector 3-4 of the LHC, a fault in the electrical bus between a dipole and a quadrupole caused mechanical damage and release of helium from the magnet cold mass into the tunnel. The design of the cryostat system of the LHC, the cooling down system of the superconductors, was one of the challenges in LHC design. Because if due to an accident, a magnet leaves its superconducting state and "quenches", a large amount of energy deposits in the magnet ($E = RI^2 - I \sim 12KA$).

A brief explanation of the 2008 incident is given in the following summary : During the ramping-up of current in the main dipole circuit at the nominal rate of 10 A/s, a resistive zone developed in less than one second to a resistive voltage of 1 V at 9 kA. The power supply which was unable to maintain the current ramp tripped off and the energy discharge switches opened to insert dump resistors into the circuit to produce a fast current discharge. So far everything worked properly, but after one second an electrical arc developed and punctured the helium enclosure and led to a release of helium into the insulation vacuum of the cryostat. After $\sim$ 4 seconds, the beam vacuum also degraded in beam pipes 2 and 1.

During the next 3 months, the 3 km-long continuous cryostat of sector 3-4 which contains 154 superconducting dipoles and many corrector magnets and quadrupole was investigated. Some of the dipoles were removed or repaired. In addition a new sensitive detection system to avoid further incidents has been installed [31, 32].

2.2.3 2010 Data taking

14 months after the incident on 20th November 2009, LHC was ready to start taking data. The beams were injected and tested on 21st and 22nd of November and the first pilot collision was on 23rd with only one bunch per beam and 900GeV of energy. During 2009, the number of bunches per beam was increased to 4 and the energy could set a new record of 2.36TeV.

After the 2008 incident, it was agreed among LHC and all detectors that an achievable energy for the first years would be less than 14TeV which is the design energy. In 2010 and after the 2009 pilot test, LHC decided not to take any risk and work at an even lower energy. After a discussion 7TeV was agreed on and on 30th of March LHC

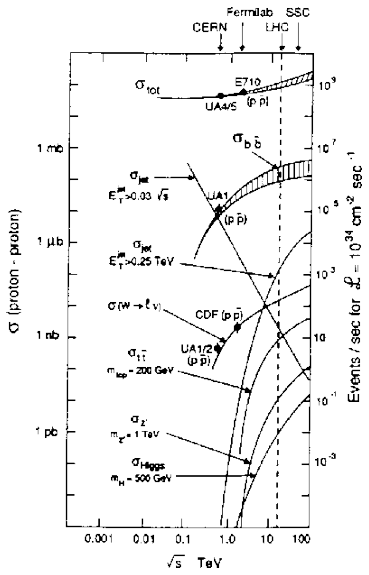


Figure 10: Cross section of important events in proton-proton collisions vs. the center of mass energy of the collision [33]

could produce the first 7TeV collision.

Reducing the center of mass energy affects the discovery power of the machine very much. In the interactions of protons with lower energies, like in Tevatron, quarks are more dominant, while increasing the energy will increase the importance of gluons in the interactions. Hence changing the energy could affect the nature of the interacting partons and so the physics could also change a lot. In Figure 10 on page 31 the cross section of some important processes are shown vs. the energy of the center of mass of the collision. Figure 11 on page 32 also compares the Higgs discovery potential of CMS experiment for different energy scenarios.

After reaching the desired energy, LHC started to increase the luminosity. As all the 7TeV collisions are comparable, all the detectors could record and analyze the data during the luminosity increase. The number of protons per bunch and the value of the emittance were easily reached to their nominal values soon after the startup. According to (29), there are two other important parameters to optimize : f, the frequency of collisions which is directly proportional to the number of bunches per beams and β^*.

As it has been already mentioned in 2.2.1, the nominal value for the number of bunches per beam is 2808 and LHC started with only 1 bunch per beam in 2009. During 8 month of continuous running in 2010, LHC could gradually increase this value

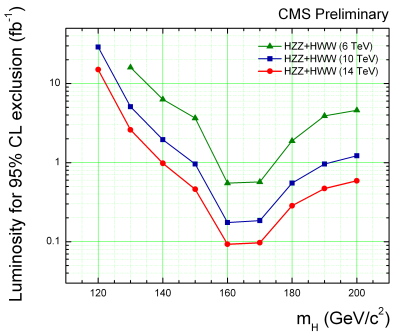


Figure 11: The needed amount of data to discover Higgs in different energies : 14, 10 and 6 TeV

to 368 among which only 348 were participating in collision.

The amplitude function at the interaction point (β^*) was also decreased from 7m in 2009 to 1.5m at the end of 2010 run. It is still far from the nominal value of 0.5m.

All of these changes and improvements allowed to raise the value of luminosity from 2.5×10^{32} in the beginning of 2010 to 1.26×10^{33} at the end of 2010 run. About $35\,pb^{-1}$ of data were collected by the main detectors, see Figure 12 on page 33.

In 2011 LHC continues to take 7TeV data and on 21st of April it could set new world record for luminosity at a hadron collider with $4.67 \times 10^{32} cm^{-2} s^{-1}$. This exceeds the previous world record of $4.024 \times 10^{32} cm^{-2} s^{-1}$, which was set by the US Fermi National Accelerator Laboratory's Tevatron collider in 2010. The new record was rapidly broken and re-broken in the following days [34].

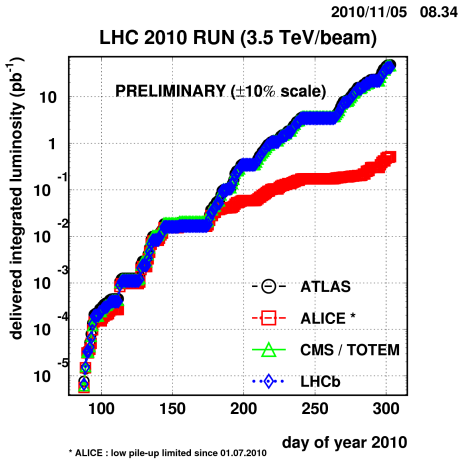


Figure 12: Total integrated luminosity collected by different experiments during 2010

3 The Compact Muon Solenoid (CMS) experiment

LHC collides the beams in 4 different points along the beam path and 4 big detectors have been installed around these collision points : CMS (Compact Muon Solenoid) at point 5 and ATLAS (A Toroidal LHC Apparatus) at point 1 which are two multipurpose experiments. LHCb in point 8 which is specialized for b-physics to measure the parameters of CP violation in the interactions of b-hadrons. ALICE (A Large Ion Collider Experiment) has been built in the fourth intersection point to study heavy ion collisions.

There are 2 more experiments, TOTEM and LHCf which are installed far from the interaction points to study the forward physics and diffraction. In 2010 the seventh detector was also approved : MoEDAL (Monopole and Exotics Detector At the LHC) which is going to be installed at the heart of LHCb and will search for the massive stable (or pseudo-stable) particles like magnetic monopoles[35] . A schematic view of the LHC and the position of different detectors is shown in 13.

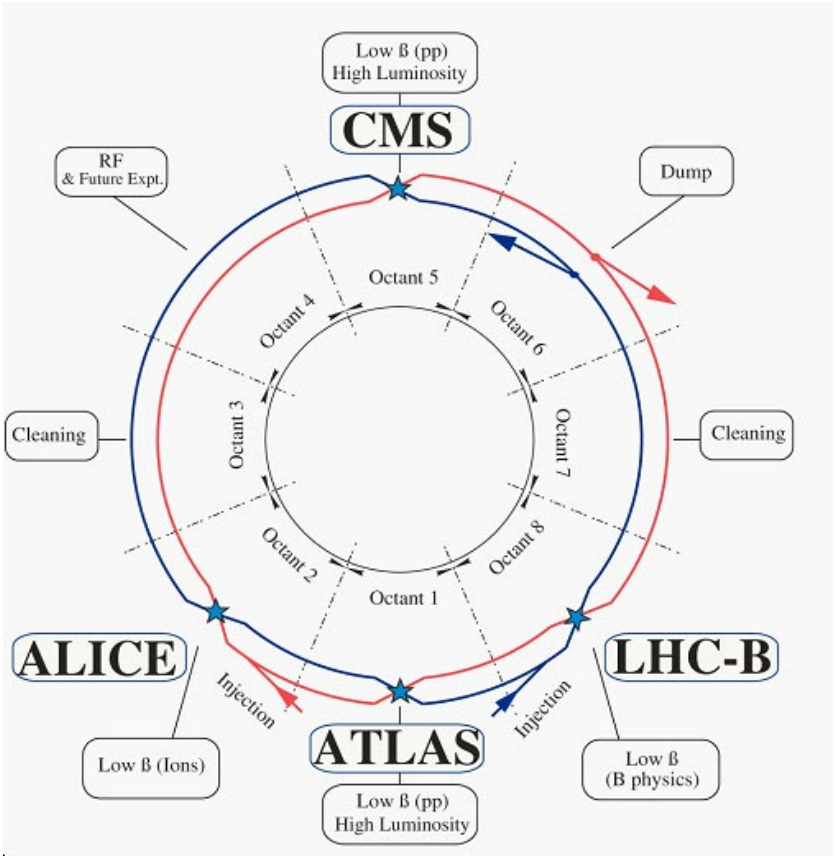


Figure 13: LHC and its detectors

The requirements for the multipurpose experiments to fulfill the LHC physics goals can be summarized as follows [36]:

- A very good Muon system to perform precise momentum measurements and identification. The dimuon mass resolution should be of the order of 1% at $100GeV$.
- Good charged particle momentum resolution and also efficient vertex reconstruction ability to tag τ and b jets.
- Precise electromagnetic energy resolution to measure the diphoton and dielectron mass as accurately as 1% at $100GeV$.
- Good missing transverse energy and dijet mass resolution which requires a hadron calorimeter with large hermetic coverage and fine lateral segmentation.

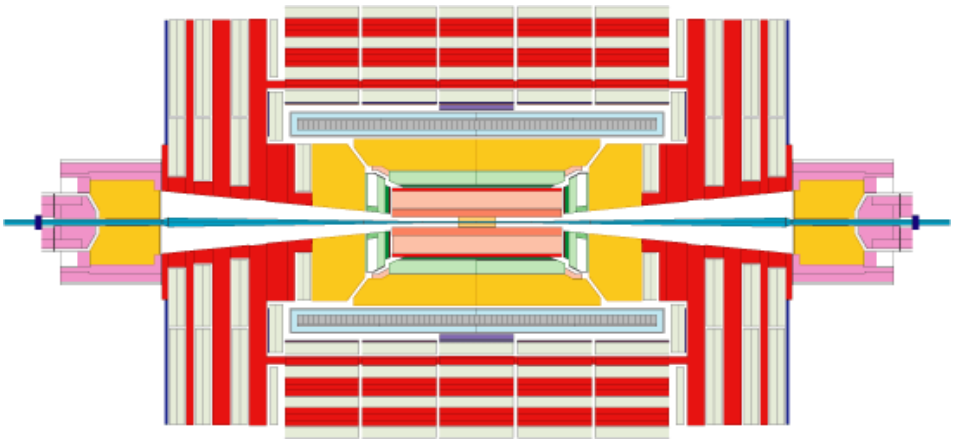


Figure 14: Longitudinal view of CMS Detector

A longitudinal view of the CMS detector can be seen in 14. As it is shown, the CMS is a cylindrical detector and the beams enter it from the two ends and the collision happens at its center.

The magnetic field is one of the key parts of each detector. Magnetic field is needed to measure the momentum of charged particles. For a massless charged particle traveling in a magnetic field we have :

$$p_t = 0.3 \times B \times r \tag{35}$$

Where p_t (in GeV/c) is the orthogonal component of the momentum of the particle to the magnetic field direction (B) and r (in meter) is the radius of the bending path of the particle. The tracking algorithms, instead of r, measure another quantity named *sagitta (s)*. As shown in 15, the relations between r and s are as follows :

$$sin\frac{\theta}{2} \approx \frac{\theta}{2} \approx \frac{L}{2r}$$
$$s = r - r\,cos\frac{\theta}{2} \approx r(\frac{1}{2}\frac{\theta^2}{4}) = \frac{L^2}{8r}$$

Replacing r with s in (35) the relation between p_t and r is found :

$$p_t = \frac{0.3 \times B \times L^2}{8 \times s}$$

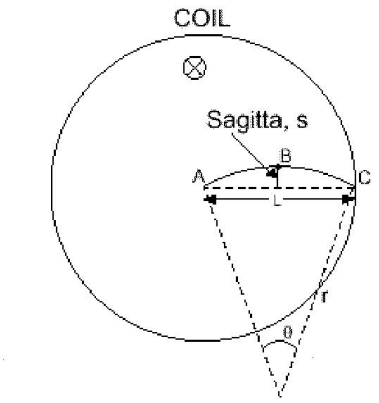


Figure 15: Schematic view of bending of a charged particle in magnetic field

For fixed values of B and L, the relative error of the p_t can be written as :

$$\frac{\sigma_{p_t}}{p_t} = \frac{\sigma_s}{s} = \sigma_s \frac{8p_t}{0.3BL^2} \tag{36}$$

The simple result is that for two different detectors, to reach the same precision on the p_t measurement, both detectors should have equal BL^2 values. Hence a compact detector like CMS needs a high magnetic field while the larger size of the detectors like ATLAS would compensate the lower magnetic field. The difference between magnetic field of CMS and ATLAS is not limited to their strength. CMS uses a solenoid with superconductors to produce the 4 Tesla magnetic field while the 2 Tesla magnetic field in ATLAS is produced by a Toroid.

CMS is made up of several sub detectors. Inside the solenoid, there are trackers, the electromagnetic calorimeter (ECal) and the hadronic calorimeter (HCal). Beyond the solenoid, the Muon system is installed. Different sub detectors of the CMS experiment are shown in 16.

A cross sectional view and also a slice of CMS detector are shown in 17 and 18 [36]. The Z-axis of the cylindrical coordinate system adopted by CMS is along the beam direction toward the Jura mountains from LHC point 5. The azimuthal angle, ϕ, is measured form the x-axis which is pointing radially inward toward the center of the LHC. Instead of the polar angleθ, pseudo rapidity ($\eta = -ln\, tan\, \frac{\theta}{2}$) is used. It can be shown easily that for massless particles, $\Delta\eta$ is invariant under the Lorentz boost along the z-axis.

In this chapter, each sub detector of CMS is reviewed. The 2010 CMS data taking

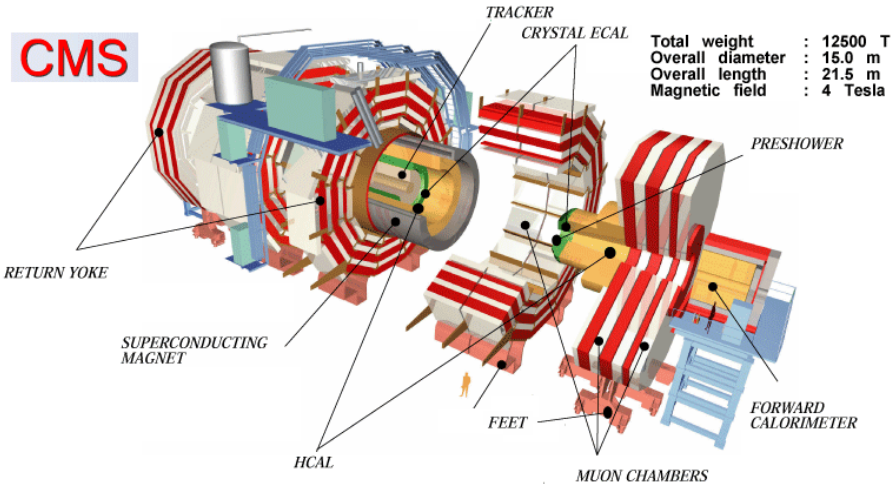


Figure 16: Different part of CMS detector in 3 dimensions

and also the efficiency of each sub detector obtained using the real data is discussed. Also the whole process of data taking including Data Acquisition (DAQ), Trigger system and Data Quality Monitoring (DQM) is explained.

3.1 The tracker

The closest part of the CMS detector to the interaction point is its tracker system. To understand the design of the tracker, the flux of charged particles in different radii of the detector should be considered. Lots of low momentum charged particles can not escape the $4Tesla$ magnetic field of the CMS detector and remain in the very first layers of the tracker. The density of charged tracks as a function of distance from the interaction point can be seen in 19. On the other hand, the information of the closest hits to the interaction point is crucial to reconstruct the vertices.

From the technology point of view, the CMS tracker can be divided into two main parts : Pixel detectors and Silicon microstrip detectors. Each part has its own design for barrel and endcap. The silicon strip detectors are divided into inner and outer detectors. A schematic overview of the CMS tracker design is shown in 20.

3.1.1 Pixel detectors

The pixel detector is made up of 3 barrel layers with 2 endcap disks on each side. The 3 barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, and have a

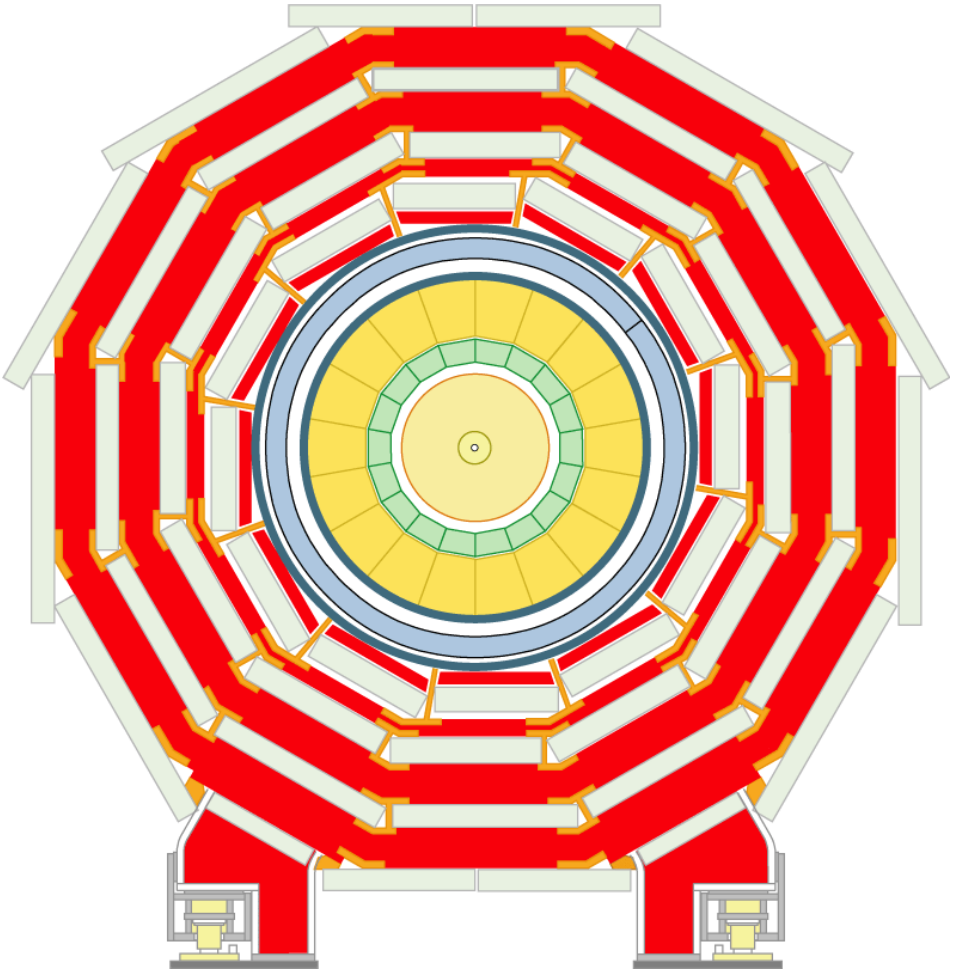


Figure 17: Cross sectional view of CMS detector

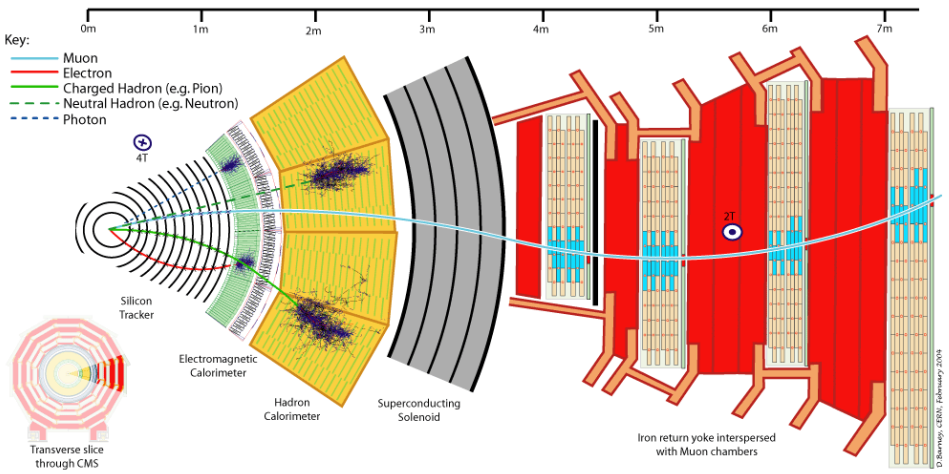


Figure 18: Slice of CMS Detector

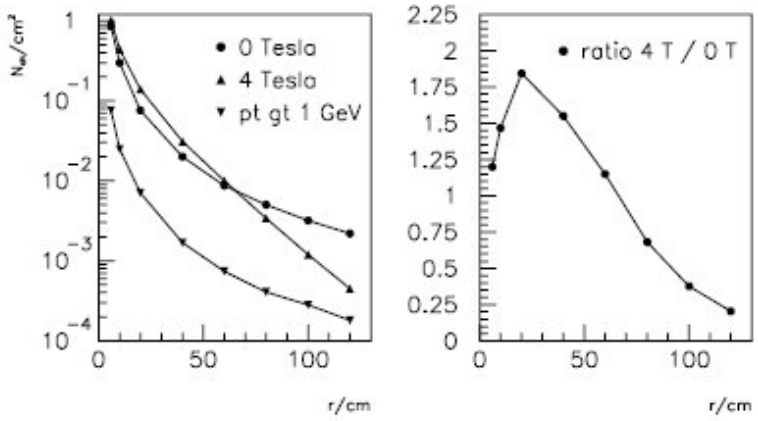


Figure 19: Charged track density as a function of distance from the interaction point [37]

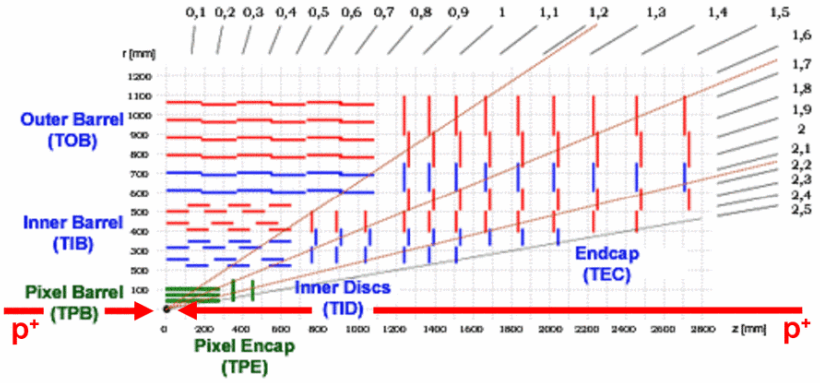


Figure 20: Schematic overview of the CMS tracker design (one quarter). As it is shown, the most inner part the pixel detectors are placed. The microstrip detectors in barrel are denoted by TOB (Tracker outer barrel) and TIB (Tracker Inner Barrel). The endcap part which extends up to $\eta < 2.5$ is divided into TEC (Tracker End Cap) and TID (Tracker Inner Disks).

length of 53 cm. The 2 end disks, extending from 6 to 15 cm in radius, are placed on each side at $|z| = 34.5$ cm and 46.5 cm. In order to achieve the optimal vertex position resolution, a design with an "almost" square pixel shape of $100 \times 150\,\mu m$ (about two hairs width) in both the (r, ϕ) and the z coordinates has been adopted. In total the pixel detector has around 65 million channels. This huge number of channels reduces the occupancy of each pixel to $\sim 10^{-4}$ per LHC crossing. This occupancy is as low as needed for an efficient tracking system.

A readout chip for each pixel exists. It splits each layer of the pixel detector into two parts : the pixel detectors and the readout chips. It can be seen in 21. A electronic silicon chip, one for each tile is attached, using an almost microscopic spot of solder using the so-called bump bonding technique, which amplifies the signal. The readout system of the tracker is described in [38].

3.1.2 Silicon microstrip detectors

After the pixel detector, the silicon microstrip detectors start from $r = 20cm$ and cover up to $r = 110cm$. Different silicon sensors for inner and outer parts and also for barrel and endcap are employed (all the different parts can be seen in 21) :

- In the first part of the barrel, so-called TIB, 4 layers of strips with thickness of $320\mu m$ and a strip pitch ranging from $80\mu m$ to $120cm$ are placed. The first two

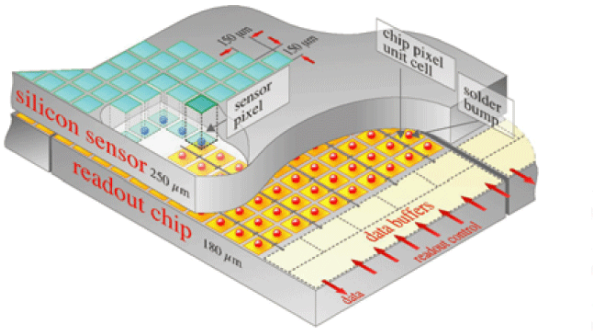


Figure 21: A layer of pixel detector

layers provide a "stereo" measurement in order to be able to measure in both $r-\phi$ and $r-z$ coordinates. The last layer of TIB is located at $r = 60cm$.

- After TIB and in the barrel, TOB with 6 layers of thicker silicon sensors ($500\mu m$) are placed. Lower values of radiation damages makes the usage of thicker sensors possible. The strip pitch in this region varies from 120 to $180\mu m$. Two layers with "stereo" modules in TOB help us to improve the position measurement.

- In the endcap and on each side, 9 disks cover the z range between 120cm and 280cm. The thickness of the 3 innermost rings of the TEC is $320\mu m$ and $500\mu m$ for the rest of it.

- TID comprises 3 small disks that fill the gap between TIB and TEC. Some of the layers in the endcap are also equipped with "stereo" modules.

The cell sizes and the granularity of the strips are so that the occupancy in the tracker remains in the order of $1-2\%$ per LHC crossing. The total area of the silicon strip detectors is $200m^2$ and it comprises 9.6 million strips [39].

The large tracker detection surface also means a large amount of material for cabling, cooling, support and electronics. In 22 the material budget is shown as a function of η, in units of radiation length and interaction length. Track finding hence needs to take possible interaction of particles within the tracker material into account.

3.1.3 Commissioning and performance with data

The calibration of the tracker has profited extensively from the hundreds of millions of cosmic rays recorded in dedicated runs in 2008 and 2009 [40]. The readout chips of the

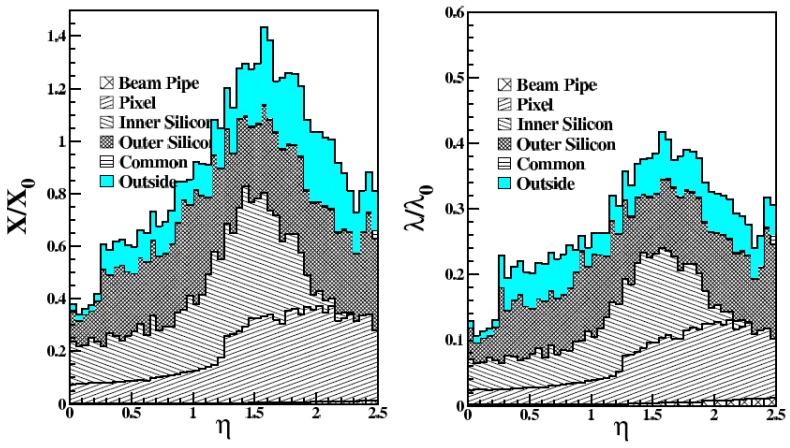


Figure 22: Tracker material budget as a function of pseudo rapidity, in units of radiation length (left) and in units of interaction length (right).

strip tracker can operate in two modes. In peak mode, used throughout 2009, a single sample is read out at the maximum of the signal pulse. While this mode has a low noise and is robust against time misalignment, the so-called deconvolution mode needs to be used to read out the tracker during LHC operations with short bunch spacing, This requires ns precision on the internal synchronization of each detector module, along with pulse shape tuning. Other calibration steps during strip commissioning include tuning of the lasers for the optical readout links, tuning of the analog baseline and the measurement of the pedestal and noise for each individual channel. The cosmic ray runs allowed for the validation and fine tuning of the above mentioned calibrations [41].

The CMS tracker registered the first LHC collisions as a remarkably well prepared detector. The performance of the tracker should be measured from different aspects. One of them is the measurement of deposited charge for high-quality tracks. It is measured for the strips and pixels with data and it agrees well with simulation as shown in 23.

Many other performance measurements and validations like "the measurement of the hit reconstruction efficiency", "pixel spatial hit resolution and correction" and also "the particle identification capabilities" are done in CMS with 2010 data and are reported by the collaboration [41, 42]. Several of them have now become routine calibration tasks, automatically run promptly on the incoming data.

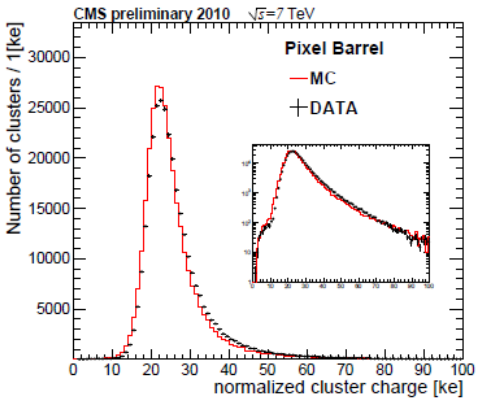


Figure 23: Normalized cluster charge in the barrel pixel for MC and data

3.2 The electromagnetic calorimeter (ECal)

The electromagnetic calorimeter surrounds the inner tracker of the CMS as shown in 16. ECal is optimized to be able to detect the di-photon decays of the Higgs boson, which is the best channel to discover Higgs with $m < 130GeV$. So it should provide an excellent di-photon mass resolution. The ECal is also split into barrel and endcaps. The barrel (EB) extends up to $\eta < 1.4442$ and the endcap part of ECal (EE) covers the η region between 1.556 and 3.0. There is a gap between $1.4442 < |\eta| < 1.556$ in the endcap which we should consider for object reconstructions and studies. A schematic view of a quarter of the ECal is shown in 24. A preshower device is placed in front of the crystal calorimeter over much of the endcap pseudo rapidity range. The active elements of this device are 2 planes of silicon strip detectors, with a pitch of 1.9 mm, which lie behind disks of lead absorber at depths of $2\,X_0$ and $3\,X_0$ [43].

The material used in ECal is crystals of lead tungstate ($PbWO_4$). About 61200 of the crystals are mounted in the barrel and 7324 in each of the 2 endcaps. These crystals have short radiation length ($X_0 = 0.89$ cm) and Moliere radius (2.2 cm), are fast (80% of the light is emitted within 25 ns) and radiation hard (up to 10 Mrad). However, the relatively low light yield ($30\,\gamma/MeV$) requires use of photodetectors with intrinsic gain that can operate in a magnetic field.

The radius of the barrel of the ECal is 129 cm and all the crystals are pointing toward the nominal interaction point. The size of the crystals is $\sim 22mm \times 22mm \times 230mm$. The front face cross section is almost equal to $\Delta\phi = \Delta\eta = 0.0174$ and the length of $23cm$ is equivalent to $\approx 25.8X_0$[36].

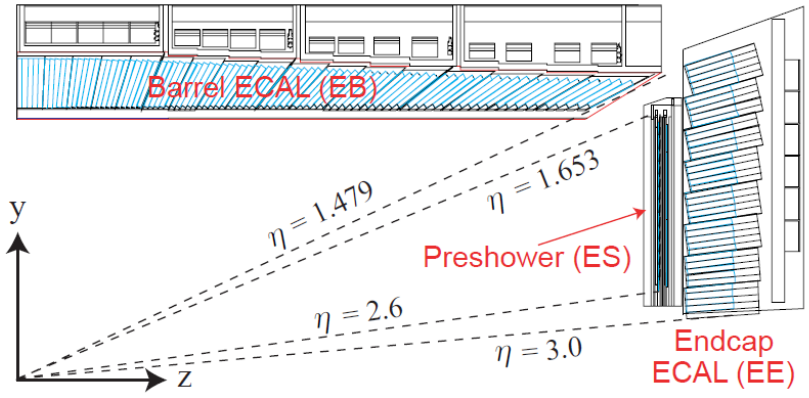


Figure 24: A schematic view of a quarter of the electromagnetic calorimeter of the CMS

The endcaps are located at $|z| = 314cm$. The crystals are installed in a $x - y$ grid instead of $\eta - \phi$. The crystals all are identical and have a front face cross section of $28.6 \times 28.6\, mm^2$ and a length of $220mm = 24.7\, X_0$.

The crystals in both endcaps and barrel are grouped in 5×5 units, known as "supercrystals" [36].

The energy resolution of the ECal can be parametrized as :

$$(\frac{\sigma_E}{E})^2 = (\frac{a}{\sqrt{E}})^2 + (\frac{\sigma_N}{E})^2 + c^2 \tag{37}$$

where the first term is the stochastic term, the second term is the noise term including both electronics noise and pile-up energy. The last term is the constant term. The value of the a and c coefficients are determined by the material used in ECal. With the lead tungstate used in the CMS electromagnetic calorimeter, $a = 2.8\%$ and $c = 0.3\%$. The value of σ_N, also obtained using the test beams, is expected to be of the order of 12%.

3.2.1 Commissioning and performance with 2010 data

For the calibration of the ECal using the collision data, there are several strategies [44]:

1. ϕ-symmetry inter-calibration is expected to provide a fast inter-calibration exploiting invariance around the beam axis of energy flow in minimum bias events. The obtained calibration factor vs. ϕ is shown in 25.

2. π^0 and η calibration uses the photon pairs selected as $\pi^0\,(\eta) \rightarrow \gamma\gamma$ candidates. At the startup, this method is also used to investigate the ECAL energy scale.

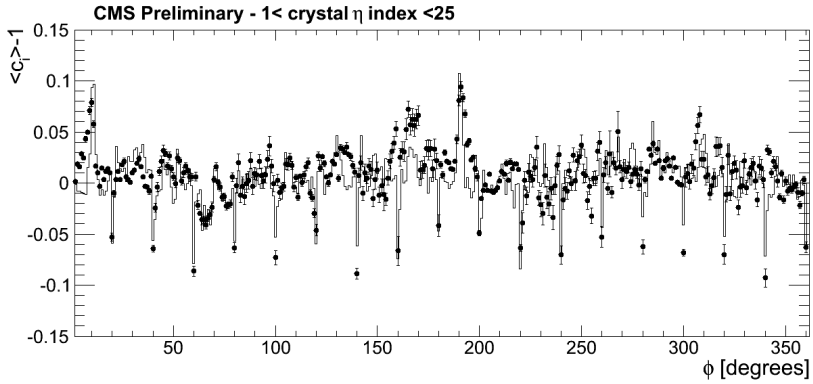


Figure 25: Average difference from unity of the inter-calibration constants derived with the ϕ symmetry method for data (solid circles) and simulation (histogram) in the crystal η index range [1, 25]. In the absence of the systematics effects, a flat distribution with statistical fluctuations around zero is expected.

3. Electrons from Z and W decays can be used to compare the energy measured in ECal to the track momentum measured in the silicon tracker. This method is expected to be the main channel-by-channel calibration tool when several fb^{-1} of data is accumulated.

3.2.2 Anomalous Energy Deposits (Spikes)

After many years of detailed studies with test beams and simulations, when LHC started to perform in 2009, some anomalous energy deposits in the ECal barrel were observed. These isolated high energy deposits which are believed to be caused by direct ionization of the avalanche photodiode (APD) sensitive volumes by highly ionizing particles - mainly protons and heavy ions. As they are usually seen in single crystals, they are often termed "spikes". The rate of these energy deposits is estimated to be approximately 1 in 10^3 minimum bias events. Studying the results of different periods of collision data taking - i.e. 900GeV, 2.36TeV and 7TeV - the rate is also observed to scale with $\sqrt{s}$ at a rate consistent with the measured increase in charged particle multiplicity [45].

The spikes could be rejected based on topological and timing characteristics. One of the good topological variables that could help us to separate the spikes is the "Swiss Cross variable". It can be defined as :

$$1 - \frac{E_4}{E_1} \tag{38}$$

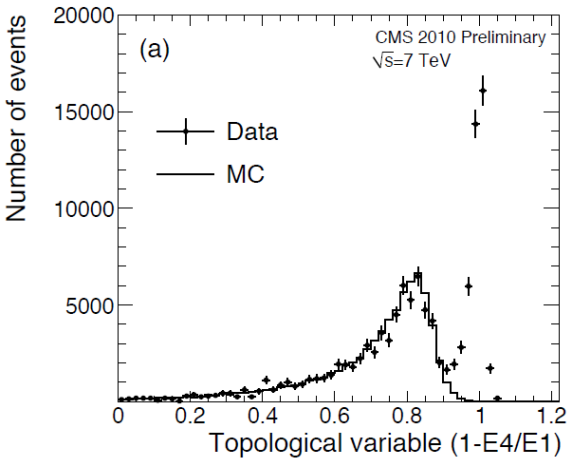


Figure 26: Distribution of the "Swiss Cross" topological variable $(1-\frac{E4}{E1})$ for the highest energy deposit in each event for data and simulation ($\sqrt{s} = 7TeV$). Only events with an energy deposit with $E_T > 3GeV$ are plotted. The two distributions are normalized to the same total number of minimum bias events, before the cut on the signal transverse energy is applied.

Where E_1 is the energy of the crystal and E_4 is the sum of all 4 adjacent crystals in η and ϕ. Computing this variable for each channel in events with total transverse energy $E_T > 3GeV$, a discrepancy is seen as an excess in data which is not present in the simulation. The comparison of MC and data can be seen in 26. Rejecting crystals with $Swiss\,Cross\,Variable\ > 0.95$ has a rejection power that depends on the signal transverse energy [46]:

- 91.8% for $E_T > 3GeV$
- 97% for $E_T > 5GeV$
- 99.2 for $E_T > 10GeV$

The timing of the signal pulse shape provides further discrimination between energy deposits from electromagnetic showers and those from direct ionization of the APD. The signal pulse shape of the energy deposits from electromagnetic showers is a convolution of the time profile of light emission from the lead tungstate crystals (80% of light emitted in 25 ns) and the response of the front-end electronics (shaping time $\tau \sim 40\,ns$). But in the signal pulse shape of the direct ionization of the APD only the electronics shaping time contributes. The comparison of the signal timing variable for all channels before

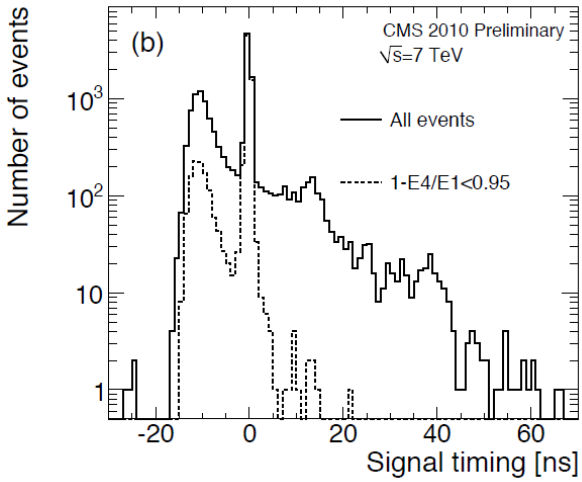


Figure 27: Reconstructed time corresponding to the maximum of the signal pulse for the highest energy deposit in each event (with $E_T > 3GeV$). The dashed histogram indicates non-isolated energy deposits that satisfy $1 - \frac{E4}{E1} < 0.95$.

and after applying the "swiss cross" cut is shown in 27. A pulse is declared 'out-of-time' if the difference between the measured and expected time is greater than 5 standard deviations [46].

The 2010 data was reconstructed several times and the cuts on these two variables were optimized during the time. The importance and effect of the spike rejection is more clear in photon and electron reconstruction which will be described in 4.1.

3.3 The hadronic calorimeter (HCal)

The calorimeter of the CMS is mostly located inside the magnet coil. The HCAL is a sampling calorimeter : to measure the energy, position and arrival time of the particles there are alternating layers of *absorber* and *scintillators*. The fluorescent "scintillator" materials produce a rapid light pulse when the particle passes through. Special optic fibers collect up this light and feed it into readout boxes where photodetectors amplify the signal. As is shown in 28, the HCal is designed in four different parts :

1. The hadron barrel (HB) part of the HCal which is made up of 32 slices each with $\Delta\eta = 0.087$ covers $-1.4 < \eta < 1.4$. Each slice consists of 72 sectors with $\Delta\phi = 0.087$. In total the HB is made up of 2304 towers. Details of the HB design,

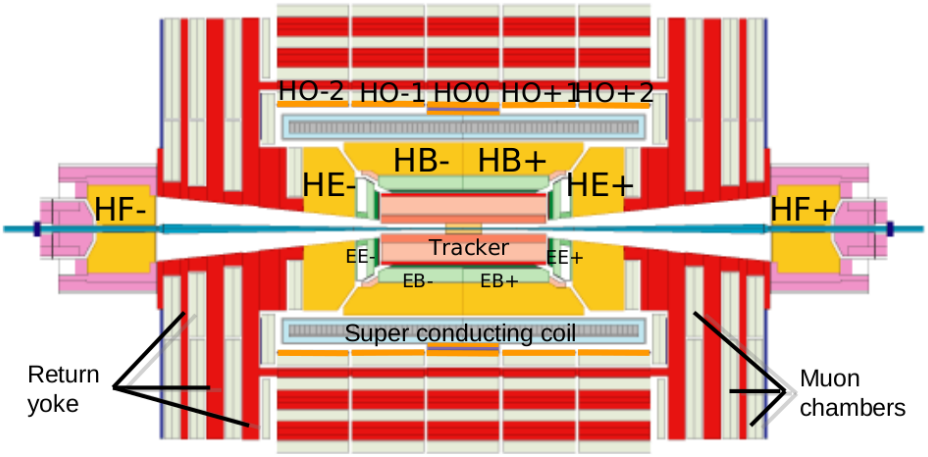


Figure 28: The position of all parts of the hadronic calorimeter in the CMS experiment

together with the performance of production modules measured in CERN test beams, may be found in [47].

2. Outside of the magnet coil and inside the barrel Muon system the hadron outer (HO) detector is placed. Divided into 5 sectors, it covers the η region between -1.26 and 1.26. The tiles are grouped in 30° sectors, matching the ϕ segmentation of the drift tube chambers. They sample the energy from penetrating hadron showers leaking through the rear of the calorimeters and so serve as a "tail-catcher" after the magnet coil. The HO also improves the E_T^{miss} resolution of the calorimeter.

3. To detect the hadronic particles in the region $1.3 < |\eta| < 3$, the HCal endcaps (HE) are designed. Each endcap consists of 14 η towers with 5° ϕ segmentation. For the 5 outermost towers (at smaller η) the ϕ segmentation is 5° and the η segmentation is 0.087. For the 8 innermost towers the ϕ segmentation is 10° , whilst the η segmentation varies from 0.09 to 0.35 at the highest η. The total number of HE towers is 2304. Details of the HE design, together with the performance of production modules measured in CERN test beams, may be found in [48].

4. The fourth part of the HCal is the hadron forward (HF) detector. It covers the $3.0 < |\eta| < 5.0$ region and its front face is located at 11.2 m from the interaction point. The depth of the absorber is 1.65 m.

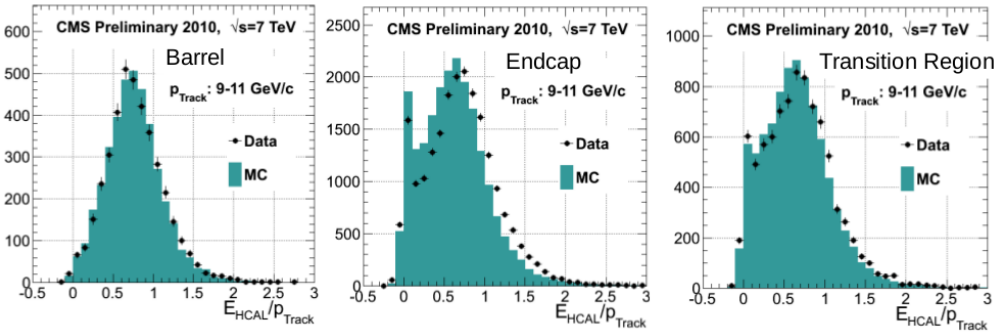


Figure 29: The HCAL response for tracks with a momentum between 9 and 11 GeV/c. HCal is divided into 3 regions :
i : Barrel ($|\eta| < 1.1$)
ii : Transition Region ($1.1 < |\eta| < 1.7$)
iii : Endcap ($1.7 < |\eta| < 2.2$)
The data is shown compared with the Geant4 based MC simulation of minimum bias events.

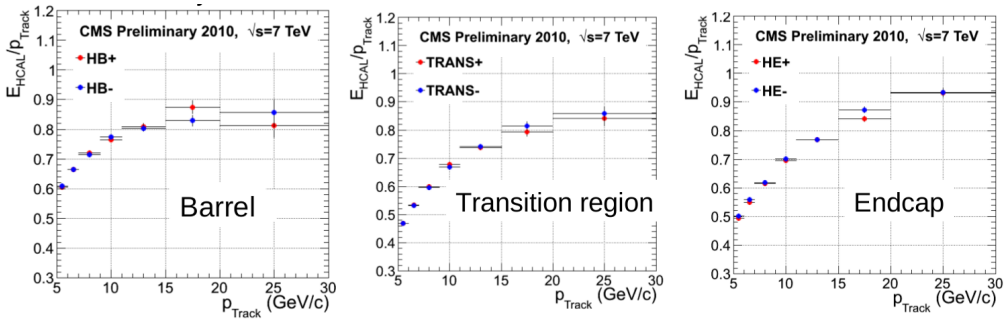


Figure 30: The mean response of the HCal, for the definition of the regions see 29

3.3.1 Commissioning and performance

To investigate the performance of the calorimeter, a method is to compare the deposited energy of a charged particle in it with the momentum measured in the tracker system. The HCAL response for tracks with a momentum between 9 and 11 GeV/c for three different regions of HCAL is shown in 29. For this study, the minimum ionizing particles (MIP) in the ECal are chosen. The mean response vs. the momentum of the track is also shown in 30 [49]. The performance was also studied with the very few 2008 beam splash events and is reported in [50].

Two classes of anomalous signals in HCAL are also observed :

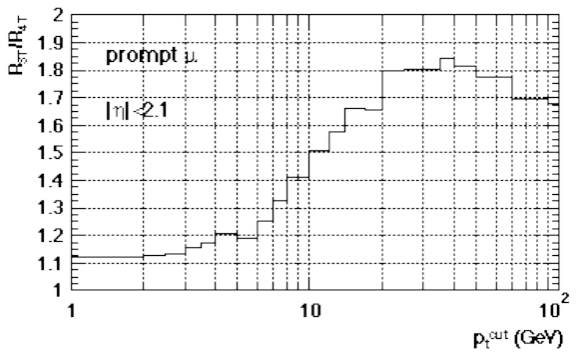


Figure 31: The ratio of trigger rates for single muon for fields of 3T and 4T

1. Electronics noise from the Hybrid Photo Diode (HPD) and Readout BoX (RBX) used for the Hadronic Barrel (HB), Outer (HO), and endcap (HE) calorimeters.

2. Cherenkov light produced by interactions in the window of the Forward Calorimeter PMTs.

Both sources of noise were studied and the needed cleaning instructions are developed and used in all object reconstructions and physics analysis [49].

3.4 The superconducting solenoid magnet

The most important part in designing a detector is the configuration and parameters of the magnetic field for the measurement of muon momenta. For CMS, the Muon system should be able to detect a narrow state decaying into Muons without any ambiguity in their signs for Muons with momentum of $\approx 1TeV/c$. This requires a momentum resolution of 10% for such high momentum Muons and hence a high magnetic field. Given the size of CMS ($overal\,diameter\ = 15m$) and $\sigma_s \sim 1mm$ and using the formula (36), the required magnetic field to achieve this precision is obtained $\approx 4T$.

There are two options for the configuration on the market, toroid and solenoid. For the CMS a solenoid is preferred for the following reasons [51]:

1. Using solenoid, the field is parallel to the beams and the bending of the muon track is in the transverse plane. The position of the interaction point (vertex) in this plane is known to a precision of 20 μm. The strong bending in the transverse plane facilitates the task of triggers based on tracks pointing back to the vertex.

Table 5: Parameters of the CMS superconducting solenoid.

Field	4T
Inner Bore	5.9m
Length	12.9m
Number of Turns	2168
Current	19.5 kA
Stored energy	2.7 GJ

Indeed the strength of the field is also important in the trigger rates. The ratio of trigger rates for single muon for fields of 3T and 4T is shown in 31.

2. In a solenoid, momentum measurement starts at $r = 0$, while for a toroid it starts after the absorber, typically at $r > 4m$. For a similar bending power the overall size of a solenoidal system is smaller than that for a toroid.

The parameters of the CMS magnet are listed in 5.

The CMS magnet has been assembled in the surface experiment hall and then moved and installed in 2006. After it and during 2 years, the data from 300 million cosmic muons were recorded with magnetic field on and at 4Tesla. The runs are known as CRAFT (Cosmic Run At Four Tesla)[40]. For safety and technical reasons, the magnetic field in all of the runs is ramped up to $3.8T$.

3.5 The Muon system

The Muon system is the last but not the least subdetector of CMS. Due to the long lifetime of the Muons ($\approx 2.2 \times 10^{-6}$) and also its small cross section while passing the matter, they play an important role in all of the recent detectors. They are charged particles and they can be traced in the tracker, but as they are almost 200 times heavier than electrons and so they are not as strongly accelerated when they encounter electromagnetic fields, and do not emit as much bremsstrahlung radiation, they can pass throughout the ECAL and HCAL without any significant energy deposit.

The Muon system of CMS is located outside the solenoid and HO [52]. It is optimized for 3 purposes :

1. To trigger on Muons

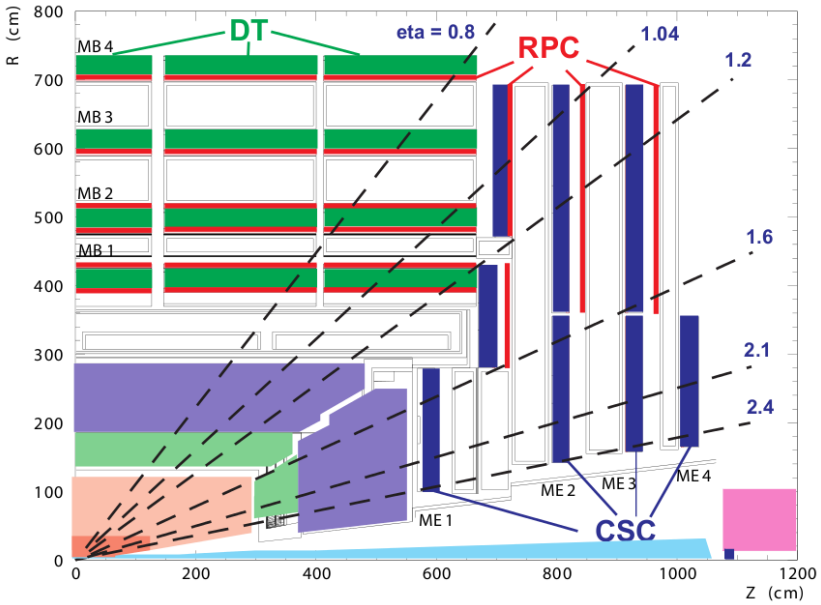


Figure 32: A quarter of the Muon system of the CMS

2. To identify the Muons

3. To measure the momentum and charge of the Muons

Similar to other subdetectors, the Muon system consists of barrel and endcap. The barrel covers the region $|\eta|$<1.2 and the endcap extends it up to $|\eta| < 2.4$ and overlaps the barrel in $0.9 < |\eta| < 1.2$. Three different technologies are used in the Muon system : Drift tubes in the barrel region, cathode strip chambers (CSC) in the endcap and resistive plate chambers (RPC) in both the barrel and endcap. A longitudinal view of a quarter of the Muon system of the CMS is shown in Figure 32 on page 53.

The rate of the Muons in the barrel is expected to be low ($< 10 particles/s.cm^2$). The return flux of the 4T magnetic field is also absorbed in the iron yoke. In such conditions drift tubes are the best choice for a precise measurement of the position of the tracks. In the endcap where the magnetic field and also the rate of particles are higher, the cathode strip chambers which are faster are more suitable. The resistive plate chambers (RPC) which provide a lower spatial resolution than the others, but provide a faster timing signal (time resolution ~ 2-3 ns) and have a different sensitivity to backgrounds are placed in barrel and endcap in order to provide trigger complementarity.

The performance of all of the different parts of the Muon system was studied during

the cosmic runs [53, 54, 55, 56, 57, 58].

3.6 The trigger and data acquisition

When LHC collides proton bunches at the highest design luminosity, around 10^9 interactions per second occur many of which are not interesting for physics purposes. Considering the network, storage and computational limitations, only ~ 100 events per second can be recorded. So CMS needs a trigger system with the rejection power of nearly 10^7.

Unlike the older design of the trigger systems, i.e. the trigger of CDF[59] and D0[60], which was designed in 3 levels, LHC multi-purpose experiments designed their trigger systems in two levels. The first level (L1) is implemented in on-detector electronics and some trigger electronics which are housed in the underground service cavern. The transmission time for signals from the front-end electronics to reach the services cavern is imposed by the size of the detector. The total time allocated for the transit and for reaching a decision to keep or discard data from a particular beam crossing is 3.2 μs. The data from the Muon system and calorimeters is processed in L1 and at this speed. During this short time, the custom hardware processors should look for "trigger primitive objects" like photons, electrons, muons and jets. Some simple global variables like the sum (vectorial or normal) of the deposited energies in the calorimeter are also considered during this step. If an interesting property or object that can pass a given criteria is seen in the event, the event will be accepted, otherwise it is rejected. The design rate of the accepted events at this level is 100KHz, i.e. one of each 10000 events can pass this step. A schematic view of the CMS trigger architecture can be seen in Figure 33 on page 55.

The events that have triggered at least once during L1 are transfered to the processor farm for online event filtering. The High-Level-Trigger (HLT) software runs on each event and reduces the Level-1 output rate of 100 kHz to a few 100 Hz for mass storage. Using the normal commercial processors for all selections beyond Level-1 maximizes the benefit obtained from the evolution of computing technology. Flexibility is maximized since there is complete freedom in the selection of the data to access, as well as in the complexity of the algorithms. HLT algorithms should follow some guidelines : Events are to be discarded as soon as possible. No need to reconstruct all possible objects, whenever possible only those objects and regions of the detector that are actually needed

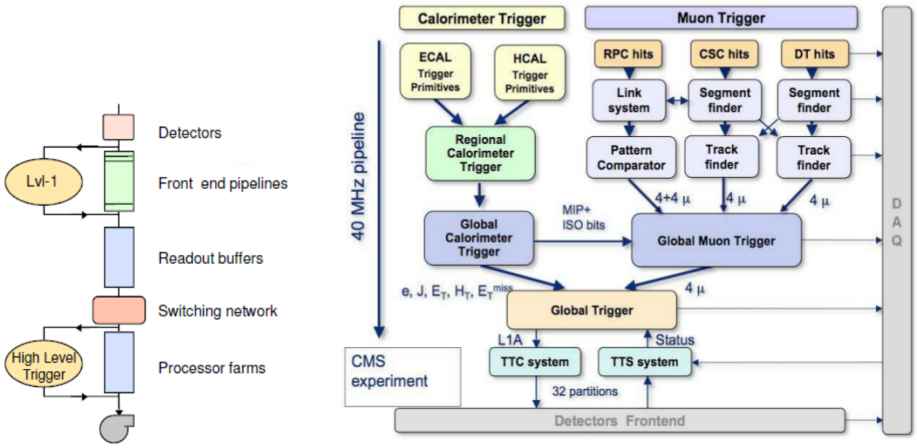


Figure 33: The CMS trigger architecture: the general scheme at left and the L1 scheme at right.

are reconstructed. In addition to the calorimeter and muon system information, the tracker pixel data and even full tracker data is accessible if necessary [61].

3.6.1 The performance and commissioning

The performance of L1 [62] and HLT [63] were studied during the cosmic runs. Efficiencies were found to be high, resolutions were found to be good, and rates as expected.

Optimizing the triggers during the runs in 2010 was one the main challenges. The LHC machince was working better than expected and the instantaneous luminosity was increased several times. The thresholds of the L1 triggers and the parameters and even algorithms in HLT are customizable. Each set of parameters and algorithms called a "trigger-menu". The trigger menu must be designed so that without loosing important data, the output rate be kept as low as expected. It was possible to predict future rates using current rates. An example is shown in Figure 34 on page 56 [64].

Another important goal of the trigger system is a proper assignment of events to the bunch crossing it originated from. Therefore, data have to be synchronized with the LHC clock every 25 ns. All CMS sub-detectors had performed the synchronization procedure during start-up of the LHC.

When an event passes the HLT, it can be categorized according the trigger which has fired. The datasets are made up of the events that fired some specific triggers. An event can be placed in more than one dataset. The physics analysis starts from the

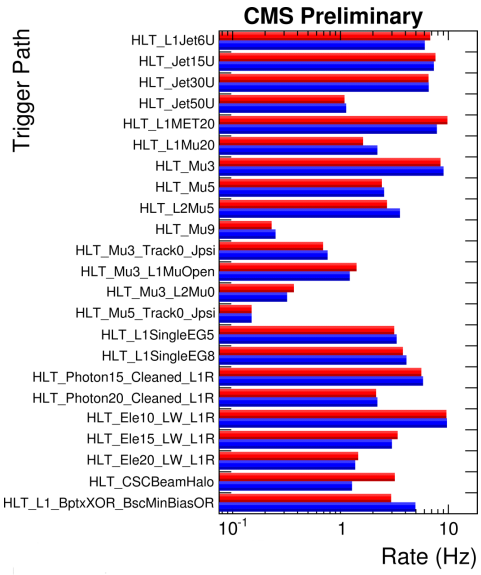


Figure 34: Trigger rates extrapolated from data. Red bins denote measured rates during data taking at the instantaneous luminosity $\mathcal{L} = 4.6 \times 10^{29} cm^{-2} s^{1}$, whereas blue bins represent extrapolated rates measured at $1 \times 10^{29} cm^{-2} s^{1}$ and scaled by a factor of 4.6.

datasets. The list of the datasets was also changed in 2010, for example at first the events that fired any of the electron or photon triggers were categorized in EGamma dataset but when the luminosity was increased they split into two electron and photon separate datasets.

3.7 CMS overall quality in 2010

The year 2010, was the first year of the LHC operation. All the objectives of the LHC machine have been achieved. A peak in the instantaneous luminosity of $2 \times 10^{32} cm^{-2} s^{-1}$ was reached in pp collisions. Since then, the amount of collected data was doubled in only a few days as shown in Figure 35 on page 57. The CMS experiment accumulated $43 pb^{-1}$ of pp data. During LHC operations CMS has obtained a high data taking efficiency of 92%. The main source of data taking inefficiency was the procedure of switching on all sub-detectors, which have to remain off for safety reasons, before stable beam conditions have been declared by the LHC's machine team. During 2010 data taking, CMS worked with event rates of the order of about 70 kHz at L1 and between 300-600 Hz output rate at the HLT.

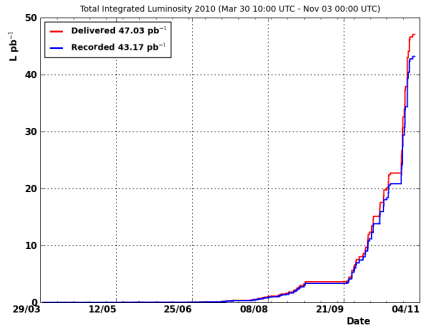


Figure 35: The CMS data taking performance in 2010. Integrated luminosity versus time delivered to (red), and recorded by CMS (blue) during pp stable beams at 7 TeV centre-of-mass energy.

3.7.1 Data quality monitoring

The procedure of monitoring the quality of data (DQM) consists of three main steps : Automatic certification and online and offline manual certification. Online and offline shifts perform visual inspection per run of a set selected histograms (shift histograms). For each subsystem the shift person takes a decision, following instructions as provided by subsystem experts, whether the subsystem is good or bad for the given run. The results from this manual assessment / visual inspection are entered in a database. At the end of the run the results are confirmed.

In the automatic certification step, certification algorithms are run in both online and reconstruction phases. The results are encoded in numbers ranging between 0.0 (bad) to 1.0 (optimal performance). Online data certification consists of DQM information only, and is presently restricted to DPG (Detector Performance Group) level monitoring.

The final offline DQM automatic certification result combines the inputs from the 3 sources of information. Appropriate certification algorithms to combine these inputs are provided by the DPG. For physics analysis only the certified data is used.

4 Physics object reconstruction

The interactions of physics objects like electrons or jets which are produced in proton-proton collisions with sub-detectors of CMS lead to production of electronic signals and pulses. Creating the event out of this information is called "reconstruction" of the event. The reconstruction starts by translating the electronic signals to physics quantities like energy and/or position. For example in the calorimeter, the signal shows how much energy is deposited. The pulses of the tracker help us to measure the precise position that the track passed through. The next step is to link these pieces of information and create tracks in the tracker or clusters, superclusters and towers in the calorimeters. Putting these together, the main physics objects can be reconstructed.

From the software point of view, the procedure of reconstruction is done by the CMS Software also known as `CMSSW`. The data structure of the reconstructed objects is very complex. A physics object may be related to one or several smaller objects. The Event Data Model (`EDM`) which has been developed by the CMS group provides a reliable framework to produce and store any kind of data. The `root` application[65] which has been developed at CERN is used as the base of `CMSSW`. The history of the events are stored in it so the time-line that each event has passed can be known, when it is created and which versions of reconstruction algorithms are used to reconstruct it.

In this section, we review the algorithms for reconstructing the objects which are used for this thesis.

4.1 Electron Reconstruction

To reconstruct each particle, we need to understand its interactions in the CMS detector. An electron first interacts with the tracker layers. The bremsstrahlung emission is an important feature of electrons that should be considered for reconstruction. It is so probable that even interacting with tracker material is enough for an electron to emit a photon. The emitted photons change the path of the electron and without any hit in the tracker reach the ECAL and make an energy deposit there. The electrons also reach the ECAL and deposit their energy there.

Putting together the position of the hits, we can reconstruct the track of the electron. The algorithms for track reconstruction are described in 4.1.1. It will be explained which algorithm is more suitable for the electron track.

In the calorimeter the crystals which contain the energy of the electron and from

the emitted photons, should be gathered together. There are again different clustering algorithms which are described in 4.1.2.

The last step to reconstruct an electron is the track and supercluster matching. There are several parameters there for the reconstruction algorithm to consider.

4.1.1 Tracking algorithms

Every charged particle interacts with the tracker while it passes through it (Pixels or strips). The position of the hits are obtained by reconstructing them. To make a track from the hits, there are two main algorithms : CTF[66] and GSF[67]. Here we describe and compare them.

CFT The first step in reconstructing a track is to generate its seed. Several hits which can be assigned to the path of a charged particle are considered as the seed. The pixel detector is capable of measuring the position of hits in the $r-z$ and $r-\phi$ planes and is the best choice to use for seed generation. At least two hits are needed to form a seed. The hits of a seed are constrained to the position of the beam and the magnetic field and so it is not as easy as choosing every combinations. In addition the matrix of errors of the parameters is calculated in this step. This step is very time and CPU consuming and it takes about 0.3 ms.

In the next step, "pattern recognition" is done. The building of the trajectory is done by the combinatorial Kalman filter method by extrapolating the trajectory of each seed using the equation of motion of a charged particle in a constant magnetic filed. In the extrapolation, the effects of multiple Columb scattering which happens in the interaction of the charged particle and the material of the tracker is considered. With the closest distance of the hits on the layers with the trajectory, a χ^2 of the track is found. It is possible that a track doesn't have any hit on a layer. It also could happen that more than one hit are found around the trajectory. These cases usually lead to more than one track.

One hit may be taken into account in more than one track, or two seeds may lead to one track. These ambiguities should be resolved. If two tracks share more than half of their hits, the one with less hits are discarded. If both tracks have the same number of hits, the track with larger χ^2 valued is removed.

After ambiguity resolution, the track fitting and smoothing is done. The parameters of the track might be biased due to the assumption used in the first steps. In order to

decrease error values, the Kalman filter algorithm is again run from the inner point of the track. After it the smoother runs from outermost hit of the track to re-evaluate the position of hits.

One of the important parameters of the track is its curvature. The transverse momentum and the charge of the track are obtained from the curvature. The impact parameter of the track (d_0) and the corresponding errors are also calculated during reconstruction.

GSF The interaction of electrons with the tracker is different from other charged particles. It is the lightest charged particle and its photon emission due to interaction with matter is not negligible. Each photon which is emitted by an electron changes the path of the electron. It makes the reconstruction of the electron track more sophisticated. To model the energy loss of the electron in each layer, the Bremsstrahlung is considered. The Bethe-Heitler [68] function includes this energy loss source too. Using the Kalman filter, some of the hits may be missed in cases where the trajectory of the track changes a lot due to emission. In the Kalman algorithm, a Gaussian distribution is used to find the hits on each layer. The GSF method uses the sum of multiple Gaussian probabilities around different central values to find the hit.

To avoid ambiguities, the maximum number of hits on each layer is required to be less than 5. The best fits are also found using the χ^2 function.

4.1.2 Clusterring algorithms in the electronmagnetic calorimeter

A photon or an electron deposits its energy in more than one crystal of the electromagnetic calorimeter. The crystals should be grouped together in order to find the exact energy of an object. A complete set of crystals which contains the energy of an object is called "supercluster". There are two algorithms to reconstruct superclusters. The Hybrid algorithm is used in the barrel. In the endcap, Multi5 $\times$ 5 algorithm first reconstructs BasicClusters and then superclusters are based on them. For electrons the supercluster is expected to contain deposits of the emitted photons too.

The position of each supercluster is calculated as a weighted average of the position of each crystal :

$$(\phi, \eta) = \frac{\sum(\phi_i, \eta_i) \times w_i}{\sum w_i}$$

The weight depends on the energy of the crystals :

$$w_i = w_0 + log\frac{E_i}{\sum E_i}$$

4.1.3 ECAL driven electron reconstruction

The raw material for the electron reconstruction is now ready. The reconstructed object are called electron candidates. Among the superclusters, the superclusters with energy more than $4GeV$ are selected. Electrons are expected to deposit all their energy in the electromagnetic calorimeter and not to reach the hadronic calorimeter. The next cut on the clusters is on $^{H}/_{E}$ which is the ratio of the hadronic energy over the electronic energy of the supercluster. This value is expected to be smaller than 0.15. To calculate the hadronic energy a cone with $\Delta R = 0.15$[1] around the electron in the HCAL is used.

To find the track of the electron, the position of the supercluster is interpolated and the expected position of the electron in the first layers of the tracker in the pixel detector is obtained. The energy of the supercluster is used for interpolation. The interpolation is done with both possible charges. If two hits are found in the pixel detector they can be used as the seed of the GSF track. After it, the GSF track algorithm starts to find the track of the electron. If such a track is found and ends in the supercluster position, they make an electron candidate together.

Momentum assignment The momentum of the track and the energy of the supercluster are two almost independent values. To produce an electron, they are asked not to be too far from each other. The tracking algorithms measure the p_T according the curvature and the radius of the low p_T tracks is smaller. So the momentum of the low p_T tracks are more precise. The formula of the error on the ECAL energy is discussed in 3.2 ((37)). To use the capabilities of the CMS detector, both measurements should be combined. The energy error depends also on the fraction of the energy of the electron which is emitted by bremsstrahlung. This fraction which is know as brem-fraction is obtained comparing the momentum of the end and beginning of the track :

$$f_{brem} = \frac{p_T^{inner} - p_T^{outer}}{p_T^{inner}} \tag{39}$$

[1] $(\Delta R)^2 = (\Delta\phi)^2 + (\Delta\eta)^2$

The electrons are categorized into three categories, according their energy and the value of f_{brem} :

- Golden electrons : $f_{brem} < 0.5$ and $^{E_{SC}}/_{P_{trk}} = 0.9$.
- High brem electrons : $f_{brem} > 0.5$ and $^{E_{SC}}/_{P_{trk}} = 0.9$.
- Cascade electrons : other electrons.

The final energy of the electron is a combination of the track momentum and supercluster energy. The weights of each one are optimized by the simulation to find the best energy resolution [69].

4.1.4 Charge assignment

The charge of the GSF track is the first source that exists for the electron charge. The electron charge is optimized using other information. As the electron charge is one of the main topics of this thesis, and to keep the consistency of the text, this optimization is explained in details in 5.3.

A 'new' variable : Skewness of hits in ϕ direction In an old study, using CMSSW_1_6 and CSA07 datasets, we tried to measure the electron charge by ECAL information to improve the measurement of charge. We developed a new variable named "skewness of ECAL hits" and investigated its correlation with charge. The results are presented in this section.

Definition To understand the charge of electron, many variables have been studied. But all of them use track information. The SuperCluster position is also used in variables like $(\phi_{innermost} - \phi_{seed})$ to improve the charge measurement [70].
Another possibility to extract electron charge information is using SuperCluster shape variables. To find such a variable, the distribution of supercluster hits in ϕ direction is studied in detail. Bremsstrahlung photons are emitted tangential to the electron path. In the presence of magnetic field, the path curvature is different for positive and negative electrons. Therefore the position of hits of emitted photons with respect to the position of electron itself, is different for positive and negative electrons in the ϕ direction.The third moment of ϕ distribution of hits can show this difference.

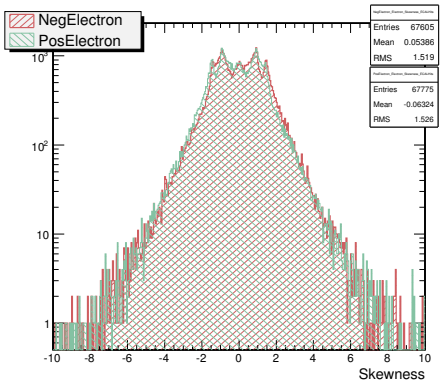


Figure 36: Skewness for positive and negative electrons separately. There is no difference between them.

The third moment of each distribution is called *Skewness* of that distribution and describes the asymmetry of that dataset from the normal distribution. Here is the exact formula for skewness of ϕ distribution of hits :

$$Skewness = \frac{\frac{1}{n}\sum_{i \in hits} w_i(\phi_i - \phi_{seed})^3}{(\frac{1}{n}\sum_{i \in hits} w_i(\phi_i - \phi_{seed})^2)^{3/2}}$$

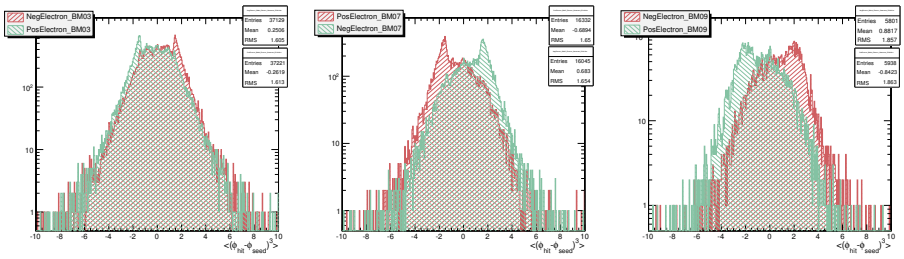
Where w_i is proportional to the logarithm of the energy of the hit and is officially used in calculating one of the electron identification variables ($\sigma_{\eta\eta}$) at CMS [71]. In Figure 36. the distribution of this new variable for positive and negative electrons is plotted.

f_{brem} and its relation to $Skewness$ As is expected from the definition of $Skewness$, it must be very dependent on the *Brem-Fraction* f_{brem} ((39)).

The dependency of $Skewness$ to f_{brem} is shown in Figures 37a , 37b and 37c.

We can combine these two variables and make the *Sharpened up Skewness* which is defined as : $exp(\frac{-1}{f_{brem}}) \times Skewnwss$ [Figure 38]

Although this variable is not correlated to other variables used for electron charge measurement, but it can not improve the charge measurement very much. It has been used for improvement and the results are reported in [72].



(a) $f_{brem} > 0.3$ (b) $f_{brem} > 0.7$ (c) $f_{brem} > 0.9$

Figure 37: *Skewness* of positive and negative electrons for different regions of f_{brem}

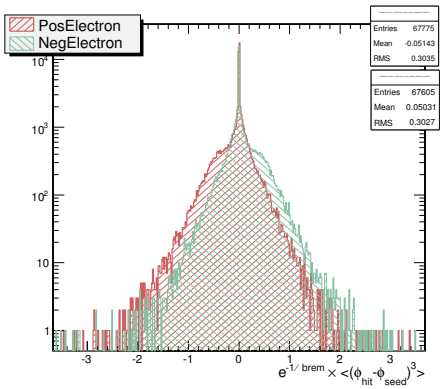


Figure 38: Sharpened up Skewness is a combination of Skewness and f_{brem} to make further studies easier

4.2 Muon

The CMS detector has a sophisticated muon system made up of RPC, CSC and DT sub systems. A powerful muon reconstruction software has been developed which reconstructs muons in the stand-alone muon system, using information from all three types of muon detectors, and links the resulting muon tracks with tracks reconstructed in the silicon tracker.

The Kalman filter technique which is described in 4.1.1 is used to reconstruct the tracks in Muon and silicon trackers separately. Muon reconstruction is performed in three stages: Local reconstruction (local pattern recognition), stand-alone reconstruction and global reconstruction. Starting from a seed, the chambers compatible with the seed are identified, and local reconstruction is performed only in these chambers. Stand-alone muon reconstruction uses only information for the muon system, while

global muon reconstruction uses also silicon tracker hits.

The first step is local reconstruction in the multi-layer chambers (DT and CSC), by associating aligned hits and builds track segments.

The stand-alone muon reconstruction uses only data from the muon detectors, without usage of the silicon tracker. Both tracking detectors (DT and CSC) and RPCs participate in the reconstruction. It starts with the track segments reconstructed in the first step as seeds of Kalman filter tracking algorithm. In the endcap, the inhomogeneous magnetic field should be considered for finding the trajectory. The position of the hits in the RPC system is not very precise and they are just used as cross check.

The third and last step of the Muon reconstruction consists of extending the muon trajectories to include hits in the silicon tracker system. The muon trajectory from its innermost stand-alone hit is extrapolated to the outer tracker surface. The effect of the interaction of muon with matter in addition to magnetic field are taken into account. It is also assumed that the muon originates from the interaction point [73].

There is another algorithm to reconstruct the muons by finding the best track that fits on the hits of silicon tracker and muon tracker together. The general idea of Tracker Muons is to reconstruct and identify muons in CMS starting from a silicon tracker track and then searching for compatible segments in the muon detectors. The energy deposition in the calorimeter can also be used for muon identification. Muons identified with this method are called "Tracker Muons" [74].

The performance of the muon reconstruction algorithm is measured using the 2010 CMS data and is reported in [75].

4.3 Particle flow event reconstruction

Instead of reconstructing different objects separately, the "particle flow" algorithm tries to reconstruct the event as a whole. The particle-flow event reconstruction aims at reconstructing and identifying all stable particles in the event, i.e., electrons, muons, photons, charged hadrons and neutral hadrons, considering all possible combinations of all CMS sub-detectors to obtain an optimal determination of their properties like direction, energy and type. The list of particles, then is used to reconstruct jets. This way, the particle content of jets are also reconstructed. The missing energy (MET) also can be reconstructed using this information. The hadronic decay of τ also is reconstructed very precisely with this algorithm.

In this section we review the reconstruction of jets and MET by the particle flow algorithm. The performance of this algorithm was studied in simulation and 2010 data. As this algorithm uses all capabilities of the CMS detector, especially its tracker, it is very efficient. The jet and MET that is used in this thesis are reconstructed by this algorithm and are denoted by pfJet and pfMET respectively.

The pfMuons are very similar to normal ones. But the particle flow algorithm is more efficient for low p_T electrons. It starts with a track instead of a supercluster and then tries to find the electron in the ECAL and reconstruct the electron. But in this analysis it is not used [76].

4.3.1 The algorithm

The main elements of particle flow objects are : tracks, calorimeter clusters and muon tracks.

Tracks play an important role though, because about two thirds of the energy of a jet is on average carried by charged particles. The momentum of charged hadrons is also measured in the tracker with a resolution vastly superior to that of the calorimeters for p_T up to several hundreds of GeV/c. To reconstruct the tracks the efficiency of the tracking algorithm should be maximized and the fake rate should be kept as low as possible. An iterative tracking algorithm is used in this step. The hits of the tracks that are reconstructed in the first iteration are removed and the algorithm is run again. The parameters of the tracking algorithm are set to be tight enough to reduce fake rate. In the next iterations, the seeding parameters become softer to increase the efficiency. Five or four iterations could usually find all of the possible tracks. The efficiency of this algorithm for muon tracks is $\approx 99.5\%$ and for charged hadrons is $\approx 90\%$. Simulation studies show that the last iterations even find the track with p_T as small as $150 MeV/c$ and an origin vertex more than 50 cm away from the beam axis and are reconstructed with a fake rate of the order of a per cent[77].

A specific clustering algorithm has been developed for the particle-flow event reconstruction, with the aim of a high detection efficiency even for low-energy particles, and towards a separation of close energy deposits. The clustering is performed separately in each sub-detector: ECAL barrel, ECAL endcap, HCAL barrel, HCAL endcap, PS first layer and PS second layer. The algorithm consists of three steps. First, “cluster seeds” are identified. They are calorimeter cells with a local maximum energy above a

given energy. Second, "topological clusters" are grown from the seeds by putting cells together with at least one side in common with a cell already in the cluster, and with an energy in excess of a given threshold[77].

Having the tracks and calorimeter clusters, they are linked together in the next step. Each particle is in one of these categories :

- Muons, with linking a track in inner tracker and muon system, and perhaps clusters with small energy in calorimeters.
- Electromagnetic charged particles, when a track matches an ECAL cluster.
- Electromagnetic neutral particles, ECAL clusters without any track.
- Hadronic charged particles, a track which ends in the hadronic calorimeter.
- Hadronic neutral particles are the HCAL clusters not reached by any track.

Many parameters exist for linking the elements and they are optimized in the particle flow algorithm.

To reconstruct the jets, any algorithm like "iterative cone" or "anti k-t" algorithms can run over the particle flow particles which are reconstructed. The particle flow objects can be given as input to the jet reconstruction algorithm[78].

Two correction are usually applied to the jet energy : absolute correction and relative correction which are also known as L1 and L2 corrections [79]. The main reason that this energy corrections are needed is the non-uniform and non-linear response of the CMS calorimeters. In addition electronics noise and additional pp interactions in the same bunch crossing (event pile-up) can lead to extra unwanted energy. The purpose of the jet energy calibration is to relate, on average, the energy measured in the detector jet to the energy of the corresponding particle jet. The correction is applied as a multiplicative factor $C(p_T, \eta)$ on each component of the raw jet momentum four-vector p_μ^{raw}.

To measure the relative correction factors, a di-jet sample is used. This method is based on the momentum conservation and assumes that in an event the total transverse momentum should be zero. So in a di-jet event we should have :

$$\vec{p}_T^{J1} = -\vec{p}_T^{J2}$$

The central region is chosen as reference because of the uniformity of the detector and jets in other parts of calorimeter are compared to the central jets. The correction

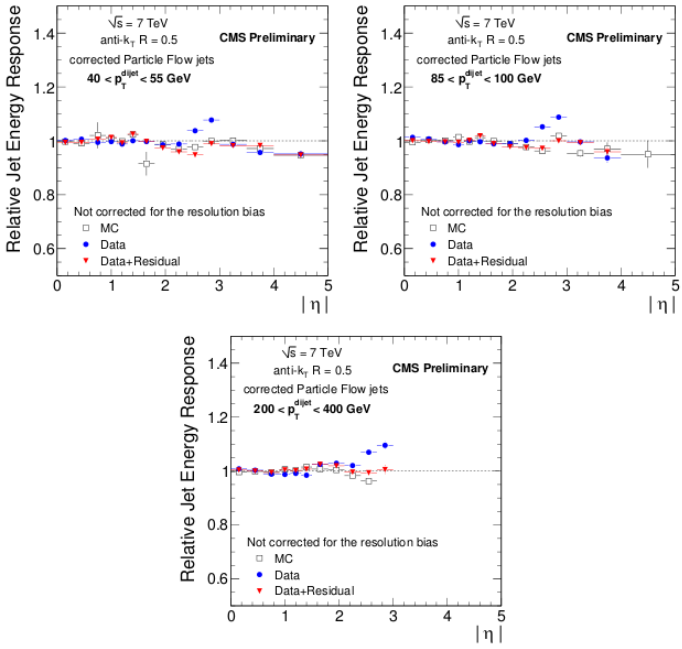


Figure 39: Relative response, R for PF jets in various p_T bins. Open squares: simulation, solid circles: data, solid triangles: data corrected with the residual calibration. [79]

factor can be obtained by this comparison by defining B factor as follows :

$$B = 2 \times \frac{p_T^{probe} - p_T^{barrel}}{p_T^{probe} + p_T^{barrel}}$$

Where the probe jet is the jet which is not tagged in the barrel. The average value of B, $< B >$, in a given η^{probe} and $p_T^{dijet} = \frac{p_T^{probe} + p_T^{barrel}}{2}$ is used to determine the relative response :

$$R = \frac{2 + < B >}{2 - < B >}$$

This factor is obtained in data and is compared with simulation in Figure 39 on page 68.

The second level of jet correction (L2) is the absolute correction. It looks in $\gamma + jet$ events and then uses the momentum conservation in the transverse plane and tries to calibrate the calorimeter towers using the ECAL energy of the photon which is more accurate. A correction factor as a function of jet p_T and η is obtained in this step[79].

The missing energy of the event can be measured by the particle flow objects easily.

It is only needed to sum the vector of transverse momentum of the pf objects :

$$MET = -\sum_{all\,pf\,objects} \vec{p}_T$$

The pfMET needs no corrections as it already includes the muons and the jet correction are applied for the jets[77].

4.4 ETH Physics Quality checking

There are several official procedures to verify and monitor the quality of the data taken in CMS. It has been described in 3.7.1. But to double check the quality of physical objects and also to verify the tools that are prepared for the analysis, a physics quality checking tool was developed by our team. The purpose of the Physics Quality Checking is to verify the quality of those variables which are actually used in the physics analysis. It is not meant to check all other parameters, which is the purpose of the on-line and off-line Quality Checking.

The Physics Quality Checking provides a set of histograms and quantities which should allow the checker to judge about the quality of a Data Set. A Data Set can be a single run or a set of runs (e.g. from reprocessing). This is defined by the user. The Data Sets are grouped and listed corresponding to given types of data. Types of data could be e.g. MC2010, Data2010, etc. A web page was made to show the results of comparisons. When a new dataset is ready, one should run the codes for checking the quality manually and then the plots will be ready to show on the web-site.

Each user can write down notes under each dataset and specify the status of the dataset. The comments of each dataset is then submitted via a html-form to the server and it saves it in a text file under the directory of the histograms of the dataset. This part of the Physics quality checking web-site was developed by me and the details can be found in B.

5 Search for super symmetry in same-sign dilepton channel

Searching for super symmetry as one of the best extensions to the standard model of particles is one of the main goals for the LHC and its detectors. The importance of this goal is considered in the design of CMS and described in 3. In this section the same-sign (SS) dilepton channel is discussed and the results of searching for SUSY in this channel in 2010 CMS data is presented.

Although this channel may not be the most powerful channel for SUSY discovery, it will be shown in 5.1.1 that due to its very specific properties and thanks to the recent developments of some kinematic variables like MT_2 and MCT, the mass of the SUSY particles that participate in the decay chain of the production of same sign dilepton can be measured via this channel. Another distinction that this channel has is the very low real background rate from standard model. Indeed, the cross section of standard model processes that result in two like sign prompt leptons is almost negligible. All of these features are also discussed in detail in this section.

This study was mainly motivated by searching for SUSY, but the results are applicable for any beyond standard model that predicts any excess in same sign dilepton channel.

In this chapter and after a brief review of the phenomenology, the event selection is shown and the yields of the 2010 data is discussed. Then in 5.3 the sources of backgrounds and the data-driven methods that we employed to estimate the backgrounds are described. And then in the last section before the conclusion, the results are used to set new limits on the new physics.

5.1 The same sign dilepton phenomenology at the LHC

In this part we continue the study of SUSY signals within the mSUGRA framework that we used in the previous sections. All of the cross sections are shown in the $m_0 - m_{1/2}$ plane, which provides a convenient way to display the signals from different sparticle production processes. For other parameters, our canonical choices are $A_0 = 0$, $tan\beta$ =2 and 10, and we adopt both signs of μ. The contours of squark and gluino masses have been shown in Figure 4 on page 21, and of slepton and chargino masses in Figure 5 on page 21.

The phenomenology of the same-sign dilepton channel has been studied by Tevatron experiments [80, 81, 82].The dominant SUSY process in hadron colliders is one of $\widetilde{g}\widetilde{g}$, $\widetilde{g}\widetilde{q}$ and $\widetilde{q}\widetilde{q}$ productions, depending on the mass spectrum of SUSY. The cross section of the SUSY production in the LHC collisions depends strongly on the mass spectra of SUSY. Figure 40 on page 72 shows the cross section for different processes as a function of the average sparticle mass[83].

We know from the MSSM lagrangian that the left handed squark can decay to a quark and an on-shell or off-shell chargino ($\tilde{q} \to \chi_1^{\pm} q'$). The virtuality of the chargino depends on the kinematics. Depending on the chargino mass, different decay chains can result in producing a lepton. If sleptons are lighter than the chargino, then the two body decays $\widetilde{\chi}_1^{\pm} \to \widetilde{l}^{\pm}\nu \to l^{\pm}\nu\widetilde{\chi}_1^0$ and $\widetilde{\chi}_1^{\pm} \to l^{\pm}\widetilde{\nu} \to l^{\pm}\nu\widetilde{\chi}_1^0$ are allowed and will dominate. Another possibility is the decay of chargino to a W boson "$\widetilde{\chi}_1^{\pm} \to W^{\pm}\widetilde{\chi}_1^0 \to l^{\pm}\nu\widetilde{\chi}_1^0$" where the W boson can either be on-shell or off-shell. So the decay chain of a squark to produce a lepton is summarized in one of these three cascades :

$$\tilde{q} \to \chi_1^{\pm} q' \to (l^{\pm}\nu) q' \to l^{\pm}\nu\widetilde{\chi}_1^0 q'$$
$$\tilde{q} \to \chi_1^{\pm} q' \to (\tilde{l}^{\pm}\nu) q' \to l^{\pm}\nu\widetilde{\chi}_1^0 q'$$

$$\tilde{q} \to \chi_1^{\pm} q' \to W^{\pm}\widetilde{\chi}_1^0 q' \to l^{\pm}\nu\widetilde{\chi}_1^0 q'$$

In a $\tilde{q}\tilde{q}$ event, if both of the squarks decay through one of these channels, two same-sign leptons are produced. The leptons are accompanied by several jets and undetectable particles that will result into missing transverse energy.

If $m_{\tilde{q}} > m_{\tilde{g}}$, almost all squarks decay to gluino and the $\tilde{q}\tilde{q}$ and $\tilde{q}\tilde{g}$ events decay immediately to $\tilde{g}\tilde{g}$. As gluinos do not couple to charginos or neutralinos directly, the only way that a gluino decays to a lepton is via an intermediate squark ($\widetilde{g} \to q\widetilde{q}$). If $m_{\widetilde{g}} < m_{\widetilde{q}}$, then the squark becomes virtual. Therefore in all $\tilde{g}\tilde{g}$ cases, each gluino decays independently to a lepton through a real or virtual squark.

Since the gluino is a Majorana fermion, it has the distinctive property of decaying with equal probability into fermions and antifermions. Thus, an excellent signature for pair production of gluinos results from events in which both gluinos decay to a chargino of the same sign, yielding like-sign dileptons (l^+l^+ or l^-l^-) in the final state. The probability for the production of like-sign and opposite-sign (OS) leptons through

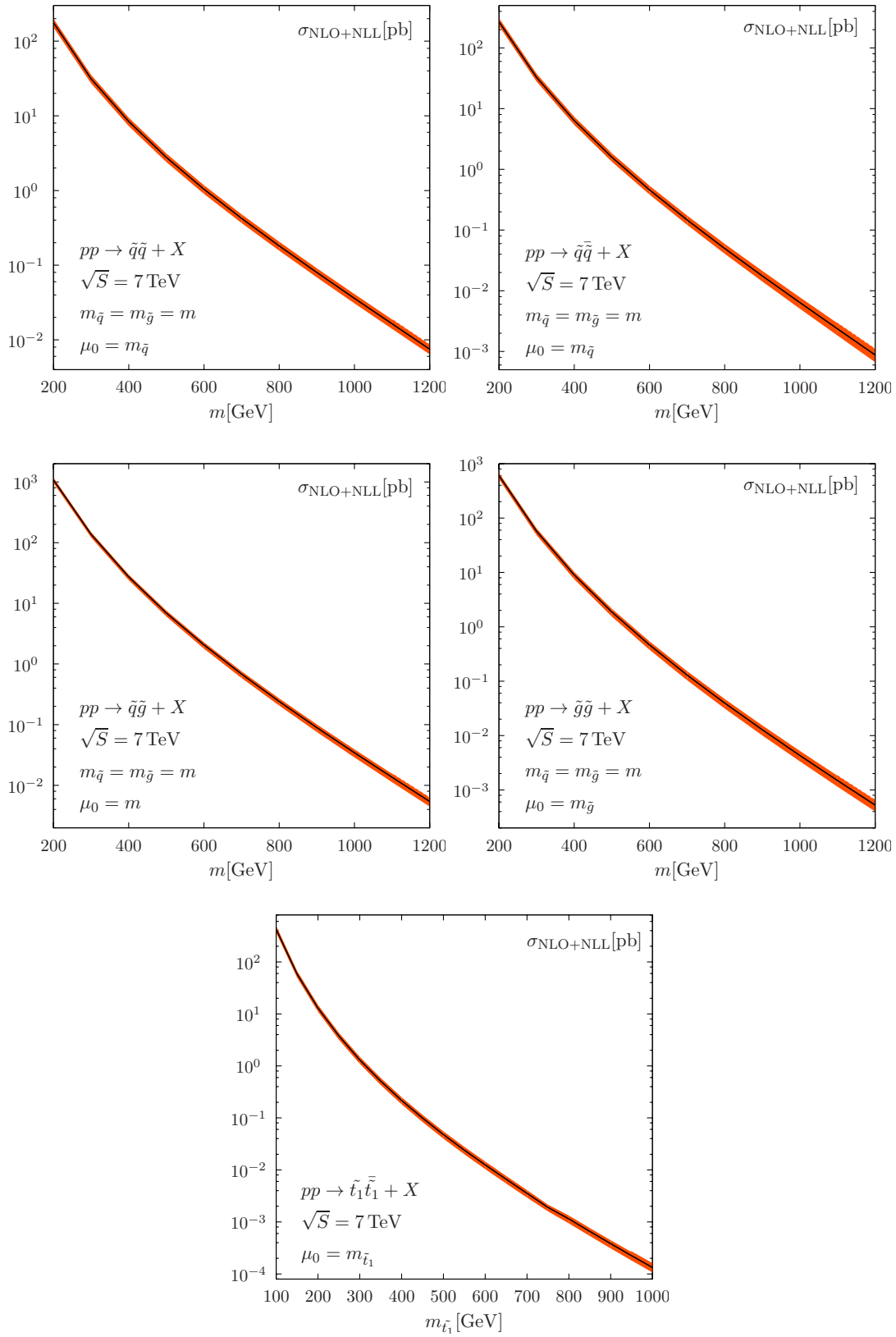


Figure 40: The NLO+NLL SUSY-QCD cross section for the individual squark and gluino pair-production processes at the LHC with 7 TeV, $pp \to \tilde{q}\tilde{q}\,, \tilde{q}\bar{\tilde{q}}\,, \tilde{q}\tilde{g}\,, \tilde{g}\tilde{g} + X$ and $pp \to \tilde{t}_1\bar{\tilde{t}}_1 + X$, as a function of the average sparticle mass m[83].

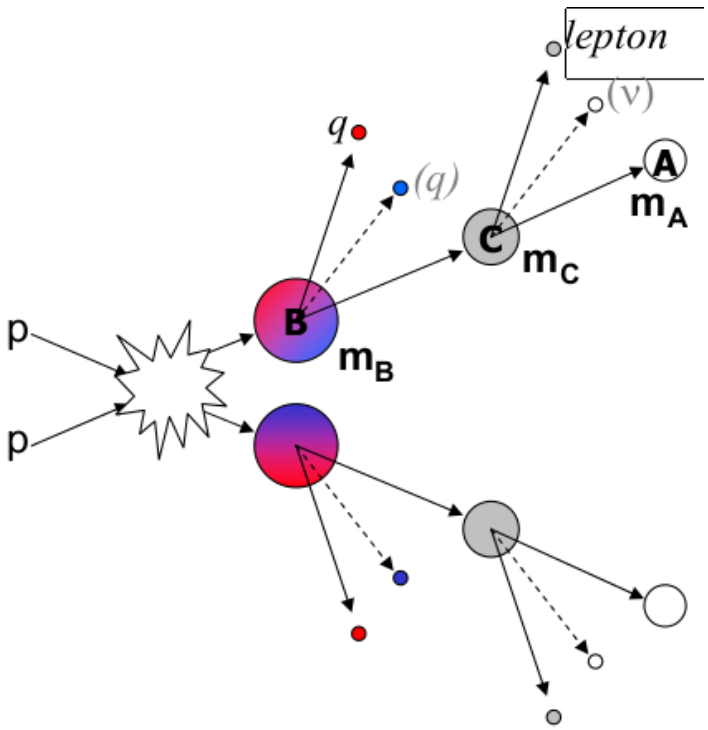


Figure 41: A schematic view to all of the main same-sign dilepton production channels:
A : LSP, dark matter motivated
B : gluino/squark
C : chargino

this decay chain are equal, and the characteristics of the two classes of final states are identical.

A $\tilde{q}$ can also decay via a χ^0 which leads to opposite-sign leptons. The production of opposite-sign leptons is expected to be larger than same-sign leptons. But this channel suffers from a much larger standard model background.

It should also be noted that the like-sign dilepton signature remains viable in models with explicit R-parity breaking in which the LSP decays and does not produce a missing-energy signature.

As has been seen, in all of the possible decay chains that result to two like sign leptons, many jets ($\geq$ 3) are also produced in association with the leptons. In the models that preserve R-Parity, the production is accompanied by a considerable source of missing energy. All of the possible diagrams that lead to same-sign dilepton events are summarized in Figure 41 on page 73.

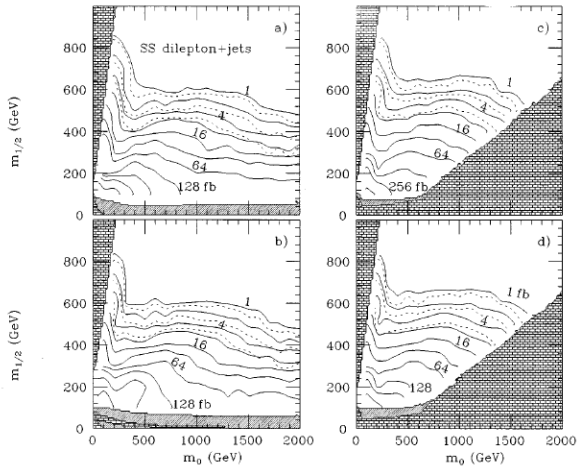


Figure 42: Contours of cross section in fb after cuts for SS dilepton+jets+MET events. The solid contours have MET and $jet_{pt} > 100$ GeV, while the dashed contours are for 1, 2, and 4 fb cross sections with 200 GeV cut. The value of $tan\beta$ is 2 for the left column and 10 for the right one. The figures in the top row is for $\mu < 0$ and the row below is for $\mu > 0$.

The cross section of the gluino pair production at LHC is given in Figure 40 on page 72. Knowing the branching ratios of the gluino decay chain, the cross section of SS dilepton production can be found. After including the leptonic branching fractions of both W bosons, one gets an overall decay of the gluino branching fraction of $\sim 1\%$ for the decay of gluino pair into a dilepton final state. Half of these are expected to be like sign since the gluino is a Majorana particle. The SS dilepton cross section is shown in Figure 42 on page 74 for different cuts on missing energy and jet transverse momentum at the generator level.

An unusual feature of Figure 42 on page 74 is the sharp kink near $m_0 \approx 400GeV$ where the slope of the contours changes. This is because of the opening up of the two-body decays of the chargino into $\widetilde{\nu}$ and $\widetilde{l}_L$. This fact can be seen in Figure 5 on page 21 in which the mass counters of sleptons and the lightest chargino are plotted. So in the left region, there are more possibilities for chargino to decay to leptons and hence the branching fraction is larger there [84, 82, 85].

5.1.1 Extracting the masses using the same-sign events

If any signal of super symmetry is observed, the next important step will be the measurement of its parameters and it starts by the masses of the super particles involved in the discovery processes. Some of the channels are more optimized for discovery while some of them can be used in the measurement phase too. It is also emphasized that the like sign dilepton events contain information about the masses of the particles that participated in the decay chains, although due to their low rate, it may not be the first channel through which the SUSY will be discovered.

Knowing the decay chain that results into the same sign dilepton events, in this section we want to see how the mass information can be extracted from these events. First we start with simpler kinematic variables to measure the mass. Then m_{T2} will be introduced as the best variable that can help us to study the like sign events.

In many of the SUSY models, the R-parity conservation and also dark-matter candidate expectation, make the diagrams and also the mass measurement too sophisticated : SUSY particles are produced in pairs and also they finally decay to a massive undetectable particle. So we should know how to deal with each of the difficulties separately.

The problem of the undetectable particles like ν, is that we can only measure the sum of the transverse momentum of all of them in the event. As explained earlier, it is called the missing transverse energy (MET) (4.3.1).

In one-branch processes that lead to undetectable particle, like $pp \to W \to l + \nu$, the transverse mass (m_T) can be used :

$$m_T^2 = (E_T^l + E_T^\nu)^2 - (\vec{P}_T^l + \vec{P}_T^\nu)^2$$

Where $\vec{P_T}$ is the transverse momentum and $E_T = |\vec{P_T}|$. In this formula the mass of the final products is assumed to be 0. It can be shown that m_T is invariant under the boost along z-direction and also it ends sharply at the value of the mass of the mother particle (w boson in this example) [86]. So the endpoint represents the mass of the intermediate particle.

There are more complicated situations in which two branches exist and both of them decay to undetectable particles. The simplest case is the full leptonic decay of $t\bar{t}$ events. A schematic view of such events is shown in Figure 43 on page 76. On each branch, Y decays to on-shell states X and v_1 with masses m_X and m_{v1} , and X then decays to on-shell states N and $v2$ with masses m_N and m_{v2} . The four-momenta of v_1 , v_2 and N

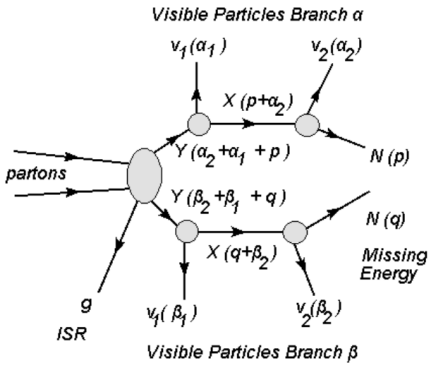


Figure 43: A schematic view of the $t\bar{t}$ events. The symbols in the parenthesis represent the value of the momentum of each particle. To map onto $t\bar{t}$ example : replace Y with top/antitop , X with W, v_1 with b-quark and v_2 with leptons. N also stands for ν.

are respectively α_1 , α_2 and p on one branch and β_1 , β_2 and q in the other branch. The missing transverse momentum is given by the transverse part of $\vec{p}+\vec{q}$. In the detectors we can measure the value of α_i and β_i and $p_T + q_T = MET$. In the $t\bar{t}$ events, our knowledge about W and ν masses allow to solve the problem of finding the top mass more easily. But in the SUSY cascade decays which is described previously, the mass of neither gluino, chargino nor the LSP is known. The longitudinal momentum of the undetectable particles can not be measured in the hadronic colliders and only the sum of their transverse momenta is measured. But a good point is that the first particles of the branches have the same mass, for example the mass of the top and antitop in the $t\bar{t}$ example or mass of gluino's or squarks in both branches of the SUSY cascade are identical. m_{T2} is a kinematic variable which can help us in such a complex situation to find the masses. Lets define it first : m_{T2} needs only the value of MET, $\alpha = \sum \alpha_i$, $\beta = \sum \beta_i$ and the mass of the particles carrying the missing energy (χ_N) :

$$m_{T2}^2(\chi_N, \alpha, \beta, MET) = \min_{p_T+q_T=MET}[max\{m_T^2(\alpha, p), m_T^2(\beta, q)\}].$$

It's more correctly to be called a 'function' rather than a 'variable', because in most of the situations in which the variable is likely to be used, the mass of the invisible object is unlikely to be known, or may only be known with a large uncertainty. So m_{T2} is normally a function of the unknown mass of the undetectable object.

This variable was introduced in 1998 in [87] and then was developed in 2003 by the same group in [88]. It can be shown easily that for each value of χ_N, the value of m_{T2}

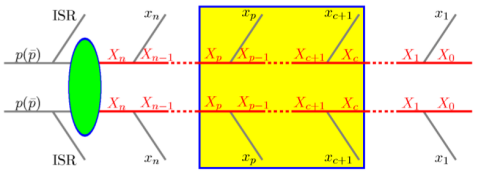


Figure 44: Each subsystem m_{T2} is parametrized by (n,p,c) where $n \geq p > c \geq 0$.

gives the smallest mass for the parent particle compatible with the event's kinematics. Thus far and thanks to the concept of m_{T2}, we could find a relation between the mass of the first and the last particle in the decay chain.

Developing this variable and defining 'subsystem m_{T2}', it has been shown in [89] that masses of all particles in SUSY-like events with two unobservable, identical particles can be measured. For an overview of the subsystem m_{T2} have a look at Figure 44 on page 77. An analytic formula for the endpoint of the subsystem m_{T2} is found in [89]. Solving the equation system of all subsystem m_{T2}'s, all the masses can be found.

The m_{T2} is used to measure the mass of the top quark in CDF experiment [90]. The problem of top quark mass has many other more precise solutions like kinematic fits and matrix methods [91]. But CDF used m_{T2} variable for the first time in a hadron collider to measure the top mass and to show how well this method works.

The usage of the m_{T2} variable in the same-sign dilepton channel is also studied in [92]. For illustration, the LM6 CMS study point in mSUGRA is considered, where the same-sign leptons most often result from chargino decays to sneutrinos. Three different techniques for determining the chargino and sneutrino masses are discussed. In all of these methods the momentum of both leptons are used and only the sum of the transverse momentum of jets as a single entity. m_{T2} is used in two of the three methods and it has been shown that the masses can be extracted with some admissible precision.

No SUSY event is observed yet, but studying the m_{T2} variable equipped us with the needed tools to start the process of the mass measurement as soon as any signal of new physics is seen.

5.2 Event selection

However many other sources of same-sign dilepton events apart from SUSY exist, in this studies we set our goal to look at the events that are more similar to what we expect from SUSY. As explained earlier in Section 5.1, same sign dilepton events from

SUSY usually contain jets and missing transverse energy. There are some theories like *Majorana Neutrinos* [93, 94] which the same-sign leptons are associated with jets, but no cut on missing energy is needed. There are also some cases in SUSY where the event with two like sign leptons does not contain as many jets as we request in our analysis. There are more new physics signatures with two like sign leptons that are not covered in this analysis, The bottom line is that this analysis is "Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy".

The sample cascade decay introduced in Figure 41 on page 73 has still many undetermined properties that are necessary to make the event selection cut flow clear. The produced hadronic energy in the event is a usual variable in event selection. It is denoted by H_T and is the scalar sum of the transverse momenta of all the jets in the event :

$$H_T = \sum_{jets\,in\,the\,event} |p_T^{jet}|$$

In the same sign SUSY cascade shown in Figure 41 on page 73, the momentum of the jets is obviously correlated with the mass difference of gluino and chargino (Δm_{BC}). So for some regions of SUSY we expect high H_T and for some other it could be lower. Comparison of the gradients of the chargino and gluino masses in Figure 4 on page 21 and Figure 5 on page 21 tells us that the higher value of $m_{1/2}$, the smaller Δm_{BC} and hence the smaller H_T. Using $35pb^{-1}$ accumulated data of CMS in 2010, we expect to be able to explore the bottom part of parameter phase space, i.e., smaller H_T values should be studied carefully.

On the other hand, the momentum that leptons carry is proportional to the LSP and chargino mass difference (Δm_{CA}). Depending on what the LSP is, the p_T of the leptons changes. So in a comprehensive analysis, low and high p_T leptons should be considered.

It has been explained in 3.6 that during 2010 data taking and as the instantaneous luminosity was being increased the HLT menu was developing and each set of runs has its own trigger menu. To start an analysis, the start point is to select a dataset which contains the data that we are looking for. The Figure 45 on page 79 represents the available relevant triggers in 2010. As is shown there, the only available HT trigger in the beginning of the run had been the HLT_HT100U which selects the events with uncorrected HT greater than 100. With increasing luminosity, the rate of this

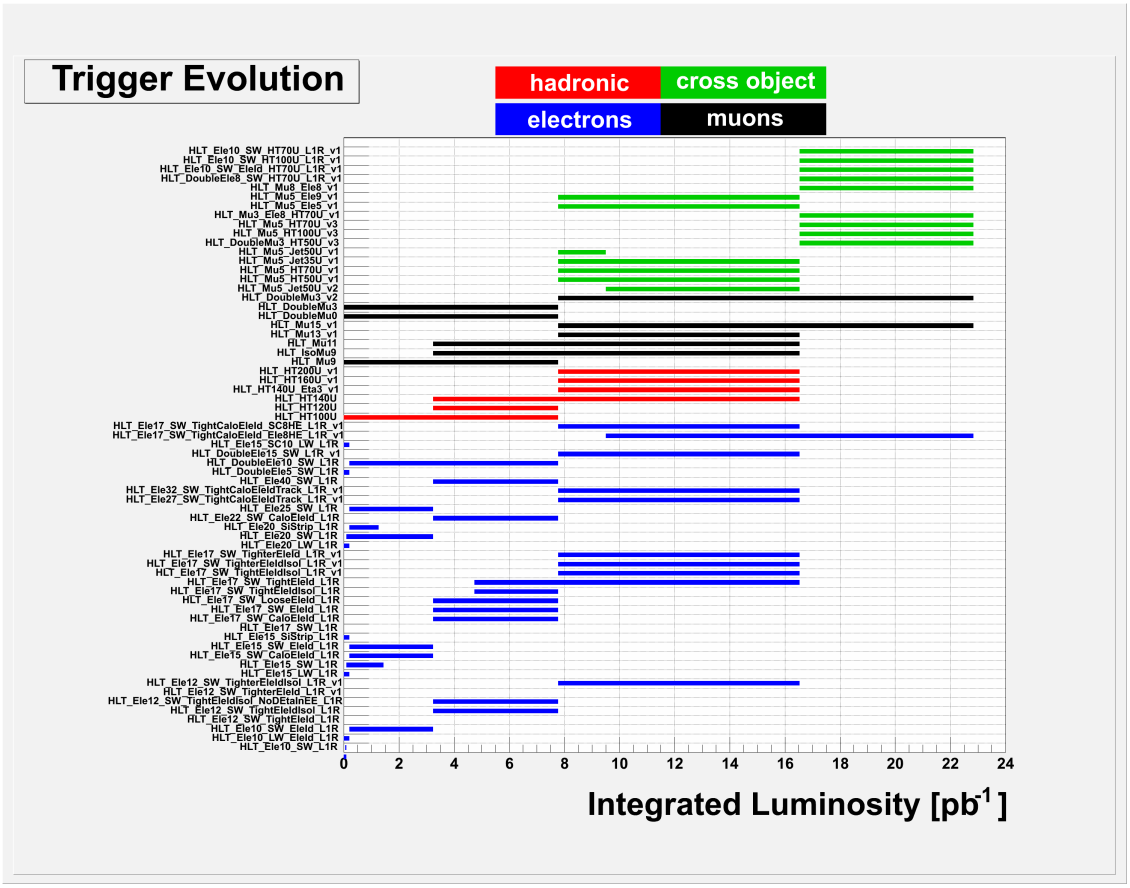


Figure 45: Trigger evolution during 2010 data taking. Each horizontal line represents the period during which a specific HLT was available in the menu. Different colors show different trigger objects. by courtesy of *Ronny Remington* who created this plot for the 2010 same-sign analysis..

HLT is increased and after a while it is dropped from the menu and HLT_HT140U, HLT_HT160U and HLT_HT200U are introduced.

The situation for the electron triggers is also the same : starting with HLT_Ele10 which accepts electrons with $p_T > 10$ and ending with HLT_Ele17. The situation for the Muon is better than for the electron.

It is obvious that in our analysis the cuts couldn't be looser than the cuts in the HLT that we use. So using the lepton triggers and datasets, we will lose low energy electrons/muons. On the other hand, using HT triggers does not allow us to select low H_T events, although the low p_T leptons are accessible in this selection.

So we decided to divide the analysis into two part :

Table 6: List of 2010 datasets which are used for this analysis.

Dataset name	integrated luminosity (pb^{-1})	Start run number	End run number
/Mu/Run2010A-Sep17ReReco_v2	3.1	131511	145762
/Mu/Run2010B-PromptReco-v2	31.3	145762	149442
/EG/Run2010A-Sep17ReReco_v2	2.8	131511	145762
/Electron/Run2010B-PromptReco-v2	31.5	137437	144114
/JetMET/Run2010A-Sep17ReReco_v2	2.9	141956	144114
/Jet/Run2010B-PromptReco-v2	31.6	141956	149442

1. Low-H_T and high lepton p_T, which is based on lepton triggers. This selection requires two leptons with transverse momenta above 20 and 10 GeV respectively, where lepton triggers are supposed to be fully efficient.

2. High-H_T and low lepton p_T, which uses H_T triggers. Softer leptons ($p_T > 5$ GeV for muons, $p_T > 10$ GeV for electrons) can be selected.

In Table 6 on page 80 the data sets which are used in this analysis are presented. All of the datasets of 'Run B' are the prompt reconstructed version while the datasets of 'Run A' are re-reconstructed versions. The CMSSW version which is used is CMSSW_3_8_2. The data has been certified on 5th November 2010 and is the most complete data set produced in 2010.

The standard recipe to clean 2010 data is applied on top of the event selection. In addition to all of the required cleanings for HCAL and ECAL which are applied in the reconstruction phase, beam scrapping events should also be discarded. A beam scrapping event is an event which has less than 25% of its tracks as high purity tracks. For the definition of a high purity track have a look at appendix A.

Each event should also have at least one good vertex. To qualify the vertices, number of degrees of freedom (ndof) should be considered instead of number of tracks associated to the vertex. Because in the vertex finding algorithms some tracks may have the weight of zero. So the number of tracks doesn't necessarily contain the information about the quality of primary vertex. A good primary vertex should has more than 4 degrees of freedom. In addition its longitudinal distance from the interaction point ($|z|$) should be less than 24 cm. If such a primary vertex does not exist in the event, the event is rejected.

5.2.1 High p_T and low H_T event selection

Triggers : As it is already mentioned, for this part of the analysis, Leptonic triggers are used. For different run intervals, different triggers were available so by firing each of them the event becomes interesting for analysis. For the Muon, the situation is simpler as always; HLT_Mu9, HLT_DoubleMu3 and HLT_Mu15_v1 are the only used muon triggers. The full list of electron triggers are as follows :

- From run #132440 to run #137028 (till the end of 2×10^{29} menu) :
  HLT_Ele10_LW_L1R, HLT_Ele15_LW_L1R, HLT_DoubleEle5_SW_L1R
- From run #138564 to run #141881 (the end of 1.6×10^{30} menu) :
  HLT_Ele10_LW_EleId_L1R, HLT_DoubleEle5_SW_L1R, HLT_Ele10_LW_EleId_L1R
- From run #141956 to run #144144 (the end of Run A) :
  HLT_Ele15_SW_L1R, HLT_Ele10_SW_EleId_L1R, HLT_Ele15_SW_CaloEleId_L1R, HLT_Ele15_SW_EleId_L1R , HLT_DoubleEle10_SW_L1R , HLT_Ele10_SW_EleId_L1R, HLT_Ele15_SW_CaloEleId_L1R, HLT_Ele15_SW_EleId_L1R
- From run #146428 to run #149442 (Run B) :
  HLT_Ele15_SW_L1R, HLT_Ele10_SW_EleId_L1R, HLT_Ele15_SW_CaloEleId_L1R, HLT_Ele15_SW_EleId_L1R, HLT_DoubleEle10_SW_L1R, HLT_Ele10_SW_EleId_L1R, HLT_Ele15_SW_CaloEleId_L1R, HLT_Ele15_SW_EleId_L1R, HLT_Ele17_SW_LooseEleId_L1R, HLT_Ele17_SW_CaloEleId_L1R, HLT_Ele17_SW_EleId_L1R, HLT_Ele17_SW_TightCaloEleId_SC8HE_L1R_v1, HLT_Ele17_SW_TightEleIdIsol_L1R_v1, HLT_DoubleEle15_SW_L1R_v1, HLT_Ele17_SW_TightCaloEleId_Ele8HE_L1R_v1

Electron selection : To make sure that an electron comes from a prompt interaction not from a jet, we need to make sure that the electron is "isolated". In order to select the isolated electrons, in a cone around the electron either in the tracker, ECAL or HCAL the sum of the transverse energy of the objects (tracks, clusters or towers) are measured. These values which are known as the "isolation values" reflect the soleness of an electron. In summing up the values, the trace of the electron itself should be discarded. It is an important point specially in the ECAL, where the emitted photons from the electron make a pattern around the electron. In the tracker the "shadow track" of the electron should not be considered. In addition to the cone size, there is another parameter to calculate the isolation value : The minimum p_T of the tracks or E_T of the calorimeter clusters. The cut on the isolation value could be a fixed number or could

be proportional to the p_T of the electron. In the second case, we divide the isolation value by the p_T of the electron and call the new variable "relative isolation". Relative isolation for the smaller values of p_T is usually larger comparing to its value for high p_T electrons. So for electron selection in this part, we decided to use the absolute isolation value for low p_T electrons ($p_T < 20GeV$) and the relative isolation for electrons with p_T larger than 20GeV. We also used the sum of the isolation values in the tracker, ECAL and HCAL. The definition of the so-called hybrid isolation which is used in this analysis can be summarized as follows :

$$EndCap\,Relative\,Iso = \frac{trackerIso + ecalIso + hcalIso}{max(p_T, 20)}$$

$$Barrel\,Relative\,Iso = \frac{trackerIso + max(0.0, (ecalIso - 1)) + hcalIso}{max(p_T, 20)}$$

In this formula the cone size of all individual isolations is 0.3 and its value should be < 0.1 for isolated electrons. The value of the ECal isolation in the barrel needs to be corrected to remove the effect of the *pedestals*.

In addition to isolation, electrons should be identified. As explained in 4.1, other charged objects like pions ($\pi^{\pm}$) or a narrow jet in which there is a charged particle, can also be reconstructed as an electron. To reject this kind of "fake electrons" as much as possible the identification cuts are needed. Standard parameters and default cuts proposed by the EGamma physics object group of CMS exist. For 2010 data, many working points for electron identification were available, each corresponding to a different efficiency of the electron selection. For example to select the electron from the $Z \to e^+e^-$ process, as two electrons are involved in the analysis, more efficiency is needed. The background rate under the Z peak is also under control. So the electroweak group of the CMS decided to use WP95 (working point 95%) from which 95% of the electron can survive. But for $W^{\pm} \to \nu e^{\pm}$ selection, using loose cuts results into large QCD backgrounds. Hence for 2010 W cross section measurement, WP85 is used [95]. In this analysis WP80 is used. A complete list of 2010 working points can be seen in [96]. The definition of the variables and the cut values for the selected working point is as follows :

- H/E, The ratio of the deposited energy in the hadron collider to the deposited energy in the electromagnetic calorimeter :
  An electron is expected to deposit all of its energy in the electromagnetic calorimeter and doesn't reach to the HCAL. So the "hadronic over electromagnetic" ratio is expected to be small. For the working point 80%, this variable should be smaller than 0.04 in the barrel and 0.025 in the endcaps.

- $\sigma_{i\eta i\eta}$, the width of the ECAL hits in the η direction
  The photons emitted from an electron along its path hit ECAL behind the electron. We studied the correlation of these hits with the electron charge in Section 4.1.4. An important feature of these hits is their narrow distribution over the η direction. It can be seen using the $\sigma_{i\eta i\eta}$ variable, the energy weighted variance of η of the ECAL hits associated to the electron, which for an electron is expected to be small.
  For WP80 electron selection, the value of $\sigma_{i\eta i\eta}$ should be smaller than 0.01 in the barrel and 0.03 in the endcaps.

- $\Delta\eta_{in}$ and $\Delta\phi_{in}$
  As explained in the electron reconstruction part, each electron is made up of two main parts : a track and a supercluster in the ECAL. The energy of the electron at the end of its path is deposited in the seed of the supercluster. So the position and the energy of the seed are comparable to the position and the energy of the end part of the track of the electron. The supercluster includes the energy of the emitted photons from the electron and so its energy and position should be compared to the information of the innermost position of the electron track. Of course the innermost position of the track should be extrapolated using the magnetic field information in order to be compared with the supercluster position.
  To identify an electron, we request the extrapolated position of the innermost position of the tracker to be very close to the supercluster position. For WP80 and in the barrel, $\Delta\eta_{in}$ should be smaller than 0.004 and the cut on $\Delta\phi_{in}$ is 0.06. The cuts for the electrons in the endcaps are 0.007 for $\Delta\eta_{in}$ and 0.03 for $\Delta\phi_{in}$.

Other criteria which are applied to select electrons are as follows :

Conversion Rejection : An important source for non-prompt electrons is the conversion of photons to a pair of electron and positron. We know that a high energy

photon converts to an electron and a positron in the presence of a heavy nucleon :

$$\gamma \rightarrow e^+ + e^-$$

The prompt photons which are produced in the proton-proton collision in LHC have enough energy and they may interact with the nucleons of the tracker material in their path to the ECAL and produce an electron-positron pair. Specially if it happens in the first layers of the tracker, the resulted electrons can fake a prompt electron. They can mostly pass the electron id and isolation cuts. So new cuts are needed to separate them.

The idea to find such electrons is to find the other product of the photon conversion, but as it is not necessarily reconstructed as an electron, we should look for it in the track collection instead of the electron collection. Hence in the CTF track collection and among the tracks which are closer than 0.3 to the electron, we look for the track of the partner. The partner track should have the opposite charge. To make sure that they are coming from a common source, the distance of the track and the electron in their closest approach should be less than 0.02. Another kinematic criteria that should be applied is to check if the partner track and the electron track are parallel in the closest approach. To verify this, the $\Delta cot(\Theta)$ of two tracks is asked to be smaller than 0.02.

To summarize, if an opposite sign track with distance < 0.02 and $\Delta cot(\Theta) < 0.02$ at its closest approach with an electron is found, the electron is assumed to be produced via a photon conversion process and is discarded from the collection of prompt electrons.

For a more detailed discussion on this topic refer to[97].

Number of missing inner hits Another possible variable by which we can reduce the number of electrons which are produced via photon conversion is the number of missing hits in the inner tracker. A photon needs to interact with matter before it converts. So the conversion happens inside the tracker material. Hence the produced electron does not hit the very first layers of the inner tracker. This means that the track of such electrons are expected to have at least one missing hit. In this study we apply the tightest cut and only accept the electrons which do not have any missing hit in the pixel tracker :

$$Number\,of\,missing\,inner\,hits = 0$$

All of these conversion rejection cuts are also part of WP80 identification cut. The cut on the number of missing hits may be looser in other analysis.

Transverse impact parameter (d_0) The distance of the track of the electron and the interaction point (beam spot) in the transverse plane shows if an electron comes from a primary interaction or is a secondary product. Mesons of the heavy flavor quarks like b and c mesons decay to electron or muon in a short time after they are created. During this short time, they fly and the decay happens a few millimeters from the interaction point. A cut on $|d_0|$ discards the leptons which are produced from the decay of these mesons. The $|d_0|$ of electrons which are used for this analysis are requested to be smaller than $0.02cm$. This cut reduces the number of the electrons from photon conversion too.

Energy of the supercluster As discussed in the electron reconstruction section, the energy of the electron is obtained from the energy of the supercluster and is corrected by the momentum of the track if needed. The track momentum is more important for low-pt electrons. So a $10GeV$ electron my have a supercluster with an energy less than $10GeV$. The energy resolution of the ECAL reduces for low energy values. So we decided to discard the electrons which are associated with superclusters their energy is less than $10GeV$.

Three charge consistency There are three different sources to measure the electron charge : GSF track charge, CTF track charge and the supercluster charge. All of these sources together with the default algorithm for the electron charge reconstruction is discussed in detail in 4.1.4. The electron charge is the charge which is common between at least two of these variables. In this thesis and to reduce the charge misidentification probability we drop the electrons when all of these charges are not in agreement. In the background estimation section, we will investigate the effect of applying this cut.

Muon selection The powerful Muon system outside the solenoid of CMS, makes the identification and measurement of the muons easier than any other object. As explained in 3.5, only muons can escape the calorimeters and exit from the solenoid magnet of

the CMS. This unique property of a muon is related to its mass and the type of the interactions it can have with matter. If a muon is produced in CMS, at first it passes through the inner tracker. As a charged particle, its track can be reconstructed using the hits it leaves in the tracker system. Then it passes through the calorimeter and deposits a small amount of energy there. Then it enters the Muon system and another track by the hits in this subdetector can be reconstructed. As discussed in the muon reconstruction section, there are two algorithms to reconstruct the muons :

- The first one uses the tracks which are reconstructed separately in the inner and muon tracker systems and tries to find the muons by matching the tracks from these two different collections. It needs to extrapolate the inner tracks to the outer point of the solenoid.
- Another method is to use an algorithm to fit the tracks globally using all of the hits in inner tracker and muon systems. The change of the magnetic field direction in the inner and outer part of CMS should be considered in this algorithm.

The muons that we use in this analysis are reconstructed by the global algorithm. In addition if their inner track is not reconstructed by the CTF tracking algorithm they are discarded. This cut rejects a lot of fake muons. Other cuts which are applied for muon selection are as follows :

1. Cuts to check the quality of the fitted track
   (a) Number of valid hits in the silicon tracker should be greater than 10
   (b) There should be at least one valid hit in the muon system (#Stand Alone Hits > 0)
   (c) Normalized χ^2 of the fit should be less than 10 ($\chi^2/ndof < 10$)
2. Cuts to discards the muons which are not produced in the primary interaction and are the result of decay of π, k or b-mesons
   (a) The first variable is the transverse impact parameter or the transverse distance of the track and the beam-spot. The importance of this variable was discussed for electrons. The same cut is also applied for the muons and the muons with $|d_0| > 0.02cm$ are discarded from our analysis.

(b) The deposited energy of a muon in the calorimeters is expected to be very low. So the energy around a muon in the calorimeter is a measure to verify if it is reconstructed correctly. It also shows if the muon is isolated or not. In this analysis we only select the muons from which the sum of deposited energies in a cone of radius 0.3 in ECAL is less than $4GeV$ and in HCAL is less than $6GeV$.

(c) Isolation is another important variable to select prompt muons. The importance and the meaning of isolation were discussed in detail in the electron selection part. We use the relative isolation for muons with p_t greater than 20 and the absolute isolation for low transverse momentum muons. The isolation formula can be written as follows :

$$Relative\,iso = \frac{trackerIso + ecalIso + hcalIso}{max(p_T, 20)}$$

The isolation is calculated in a cone of 0.3 around the muon and only the muons with isolation less than 0.1 are accepted in this analysis.

Jet selection The *particle flow jets* are used for this analysis. This particle flow algorithm has been explained in 4.3. It was also mentioned that the value of the energy of the jets needs several corrections due to the needed calibrations of the calorimeters. The p_T of the jets, after applying L1 and L2 corrections, should be greater than $30GeV$ for this analysis. As the particle flow algorithm is largely based on the tracking, the jets outside of the tracker region are discarded. The kinematic cuts can be summarized as follows :

- $p_T > 30GeV$, after applying L1 and L2 corrections.
- $|\eta| < 2.5$, to be inside the inner tracker region

Jet-Electron cross cleaning : In addition to these cuts, the jets should be cleaned with respect to the electron collection. The jet algorithm may reconstruct an electron also as a jet. It is necessary for us to cross clean the jets from the electrons to avoid any double counting. The electrons which are found after applying the cleaning and isolation cuts are assumed to be the real prompt electrons. If any jet within the cone of 0.4 around such an electron is found it is discarded.

Loose jet id : There are other criteria proposed by the jet-met PAG[2] to select clean jets. We use the *loose* set of the criteria which includes these cuts :

- The electromagnetic fraction of the jet should be greater than 1% ($emf > 0.01$). This cut is needed to reject the electrons which normally deposit most of their energy into the electromagnetic calorimeter. This cut is very correlated to the jet-electron cleaning.
- The energy of a physical jet is expected to be distributed over many calorimeter hits. N_{90hits} represents the number of calorimeter hits containing 90% of the jet energy. We require $N_{90hits} > 1$.
- The deposited energy in the hottest hybrid photodiode in the HCAL should be less than 98% of the jet energy ($f_{HPD} < 0.98$).

Missing transverse energy (MET) In this analysis, we use the MET which is reconstructed by the particle flow algorithm (pfMET). The reconstruction of the MET via this algorithm considers all of the possible objects including muons and so no further correction is needed.

Event selection After selecting clean events according the fired triggers and selecting the objects by the procedure described above, now we can select the same-sign events.

The first step is to select the events with two same-sign leptons (e or μ). The p_T of the hardest lepton should be greater than $20GeV$ and the p_T of the other lepton should be $> 10GeV$. These cuts on the transverse momentum of the leptons are relatively high, but we can not avoid it due to the lepton triggers that we use to select the events.

There may be events with more than two leptons satisfying these criteria .In each event with more than two leptons, at least a pair of same sign leptons exists. In this analysis there is no upper cut on the number of leptons and multilepton events are also accepted.

We have to make sure that the leptons are not produced from the Z boson decay or decay of any other resonances like J/Ψ and Υ. So in each event with two or more leptons, the invariant mass of the same-flavor lepton pairs is computed. Any opposite-sign pair in the Z boson mass window and the pairs with invariant mass less than $12GeV$ are dropped. This step can be done in the lepton selection level.

[2]Physics object group

In the events with two same-sign leptons, we ask for two or more selected jets. As the p_T of each selected jet is at least $30GeV$, this cut implies also a cut of $60GeV$ on the hadronic activity of the event (H_T). This cut on the H_T is low enough and is possible because of using the lepton triggers.

All of these same-sign events belong to one of these categories : either the leptons are from the same generation ($e^+e^+ - e^-e^- - \mu^+\mu^+ - \mu^-\mu^-$) or from different generations : ($mu^+e^+ - \mu^-e^-$). We apply different MET cuts for these different categories and ask a MET value greater than $30GeV$ for same flavor events and $20GeV$ for opposite flavor events.

So all of the event selection cuts can be summarized as follows :

- Cleaning the same-flavor leptons according to their invariant mass by discarding all of the leptons from Z boson decay

  $$|m_{ll} - m_z| < 15GeV$$

  or the decay of heavy mesons

  $$m_{ll} < 12$$

- Select the events which have two same-sign leptons and the highest p_T is greater than $20GeV$ and the other is greater than $10GeV$.

- Then apply a cut on the number of selected jets : $N_{jets} > 2$. This cut implies that $H_T > 60GeV$.

- MET $> 20GeV$ for the opposite-flavor events and $> 30GeV$ for the same-flavor events.

5.2.2 Events with high H_T and low p_T leptons

The second part of the analysis, as was already mentioned, is the study of the events with softer leptons and high hadronic activity. It is another existing possibility according to the trigger menus of the 2010 data taking. The procedure of the object cleaning and event selection is very similar to the previous section. The differences are mentioned here.

Triggers : The hadronic datasets (JetMET and Jet) which include the events which fired the H_T triggers are used. Here is the list of all of the triggers used in different run ranges :

- HT100U (run 140160-147116 (8 pb^{-1})),
- HT140U (run 147196 - 148058 (8 pb^{-1})),
- HT150U_v3 (runs after 148822).

The list of the H_T triggers is simpler than the leptonic triggers which was used for the high p_T lepton selection.

Electron Selection : There are two differences with respect to the electron selection in the previous section :

- There is no cut on the energy of the supercluster of the electron
- The definition of the isolation does not change, but the isolation cut is a bit looser in this analysis and is changed from 0.1 to 0.15.

$$\textit{Relative isolation within cone}\ 0.3 < 0.15$$

Muon selection : The main difference in selecting the muons in this analysis is that muons with lower transverse momentum are selected. The cut on the p_T of the muons is chosen to be $5GeV$. Such a low p_T selection is not possible for the electrons due to the high rate of QCD backgrounds for low p_T electrons.

Another difference that the muon selection has in this part of the study is the looser cut on the isolation. Like for the electrons, the cut on the muon isolation is increased to 0.15.

Jet and MET selection The jets and the MET which are used in the high H_T and low p_T lepton analysis are the same the low H_T and high p_T lepton analysis.

Event selection The first step in the event selection is to remove the leptons which are produced from the Z boson decay or the decay of a heavy meson. As low p_T leptons are participated in this analysis and to keep more events, the lower cut on the m_{ll} is $5GeV$ instead of $12GeV$.

In the events with two same-sign leptons, we ask for two or more jets. In addition a cut on the scalar sum of the transverse momenta of the jets (H_T) is also applied and only events with $H_T > 300GeV$ are accepted. The value of $300GeV$ is chosen according to the turn-on curve of the triggers to make sure that the efficiency of the triggers is 1.

The cut on MET is $30GeV$, independent of the sub-channel the event belongs to.

5.2.3 Results in 2010 data

The result of applying all of the above cuts on the 2010 data can be easily summarized as follows :

- 3 same sign opposite flavor ($e\mu$) events were found with the low H_T and high lepton p_T selection. The information of one of these events is shown in Table 7 on page 91 and the transverse view of this event is also show in Figure 46 on page 92.

- In another selection with a lower cut on momentum of the leptons and tighter cut on the H_T, just one event with two same-sign electrons could survive after applying all of the cuts. The information of this event and its display can be seen in Table 8 on page 93 and Figure 47 on page 92.

The reason why the three events with high p_T leptons did not pass the low p_T selection cuts is the low H_T of all of those events. But the answer to the question of why the only ee event of the low p_T selection is not selected in the high p_T selection is not so easy. As is seen in the Table 8 on page 93, the momentum of the electrons are high enough (75.8 and 20) to pass the first selection cuts. But their problem lies in the isolation of one of the electrons which is between 0.1 and 0.15.

Table 7: Properties for one of the $e\mu$ events found in the data using the high-p_T selection cuts. All of the numbers are in GeV units.

H_T	pfMET	N_{jets}	ep_T	μp_T	$Jet^0 p_T$	$Jet^1 p_T$	$Jet^2 p_T$
157.4	29.5	3	54.3	27.6	88.0	69.4	33.4

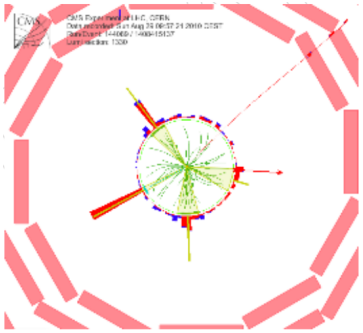


Figure 46: One of the eμ events found in the data using high-p_T selection cuts.

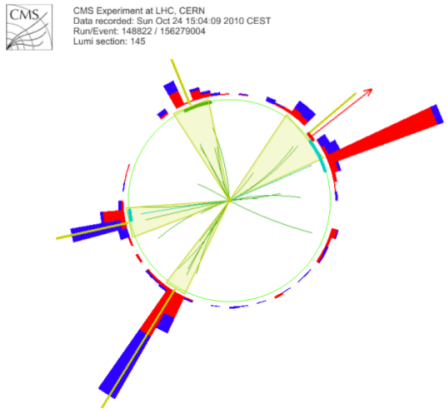


Figure 47: *ee* event found in the data using the low-p_T selection cuts

5.3 Background processes

The important question arosen by the observation of four like-sign events in 2010 data is wheater they are signals of new physics or are explainable by the standard model of particles. In addition to the SUSY processes which result into the same-sign dilepton production and have been discussed in detail in Section 5.1, three other sources for such kind of events :

- There are processes in the standard model of particles which lead to two like sign leptons from proton-proton collision. These processes, in which two gauge bosons are produced, have a very low cross section and are negligible for 35 fb^{-1} of data.

- The strong magnetic field and the very precise inner tracker of CMS increase the precision of the lepton charge reconstruction however there could be still errors on it. Specially for the electron, which should be reconstructed in a more crowded environment and the bremsstrahlung emissions which makes the path of the electrons more complicated. In an opposite-sign dilepton event, if the charge of one of the leptons is wrongly reconstructed, the resulting same-sign event may pass other cuts of our event selection and counts as a background to our analysis. The effect of full leptonic decays of $t\bar{t}$ and opposite sign WW events should be considered as the main source of opposite-sign events in our event selection region. Thanks to the high precision of the Muon system of CMS, The charge mis-identification rate for Muons is negligible.

- Another background which appears as the main one is due to misreconstruction of the leptons i.e. selected lepton candidates which are not primary products of the the main interaction. Jets often produce leptons that are wrongly reconstructed as primary. In the rest of this document, we refer to these leptons as *fake leptons.* A fake lepton may originate from a real and isolated lepton within a jet, e.g a leptonic b decay, or mis-reconstruction of a narrow jet or a $\pi^{\pm}$ inside a jet.
  Measuring the number of backgrounds from the fake leptons is more complicated

Table 8: Properties of the *ee* event found in the data using the low-p_T selection cuts. All of the numbers are in GeV unit.

H_T	pfMET	N_{jets}	e^0p_T	e^1p_T	Jet^0p_T	Jet^1p_T	Jet^2p_T	Jet^3p_T
343.3	45.6	4	75.8	20.0	160.1	85.3	58.8	39.2

: we first need to develop a method to measure the probability that a non-lepton object fakes a lepton and then try to estimate the number of backgrounds due to fake leptons.

Another complexity is that there are different background configurations : Some events like $W \to e\nu$ or semi-leptonic decays of $t\bar{t}$ have one real lepton and with only one fake lepton they are reconstructed as dilepton events. If the charge of the fake lepton is the same as the charge of the real one, which depends on the origin of the fake lepton, the same-sign event may pass our event selection. But in a QCD event in which no real lepton can be found, if two fake leptons are reconstructed, there is a chance for them to be like-sign.

In the rest of this section, we will try to estimate the number of events from all of the three sources which are mentioned above to find out if the 4 observed events in 2010 by CMS experiment originate from new physics or not.

5.3.1 Same sign dilepton events from standard model

Very rare sources of two real, prompt and same-sign dilepton events exist in the standard model of particles. These events are not yet observed in the proton-proton collisions directly due to their very low cross sections and we can only study them through the MC generators. Some of them have been observed in Tevatron experiments and it helps us to have a more accurate estimation.

Any event with more than two prompt leptons can be considered as a background for our study, as at least a pair with the same-sign exist in them. In most of the cases, a Z boson is produced in multilepton events. Consider a ZW event as an example :

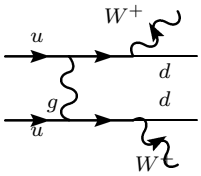
$$pp \to ZW \to (l\nu)Z \to (l\nu)(ll)$$

In addition to the very low cross section of such a process, the cut we have applied on the mass of all of the same-flavor lepton pairs to discard all of the leptons from Z decay, does not let such events remain in our selection. These events do not have a significant hadron activity and so the H_T cut and the cut on the number of jets would reject them too.

ZZ events are another type of events which include Z bosons and produce more than two leptons. Considering all of the cases, even the τ decays to electrons an muons, the rate is still negligible and the applied cuts would reject all of them.

The events in which no Z boson is involved could also be interesting. Apart from the production of a W boson associated with $t\bar{t}$ pair and three W production which have not been studied yet in proton-proton collisions, two interesting sources for same sign W pairs exist :

- $qq \to q'q'W^{\pm}W^{\pm}$
  Two bremsstrahlung W bosons could be emitted and two new quarks in addition to the same-sign pair of the $W^{\pm}W^{\pm}$ produced.



  The cross section is different for different charges. For two negative W's it is lower, similar to the cross section of a sole W^-which is lower than the cross section of the W^+ production. The cross section of $W^-W^-$is estimated to be $\approx 64fb$ and this value for W^+W^+ is around $188fb$ [98, 99, 100].
  It means that in 2010 data of CMS and within $35p^{-1}$ of accumulated data, ≈ 9 events from this process are expected. Considering the branching fraction of the $WW \to (l\nu)(l\nu)$ which is $\approx (2/9)^2 = 5\%$, less than half an event is expected to be produced via this channel. Applying all of the other cuts on the jet and MET in the MC confirms that the effect of this source of real same-sign dilepton production is insignificant.

- Double parton interaction and $W^{\pm}W^{\pm}$ production $(2 \times (q\bar{q}\prime \to W^{\pm}))$
  The cross section of this process is around $200fb$ which is less than the other mechanism of $W^{\pm}W^{\pm}$ production [101]. The number of jets produced within this process is also lower. Hence the contribution from this background to our study is even more negligible.

5.3.2 Charge mis-identification

As mentioned earlier, mis-identifying the charge of electrons in an opposite charge dilepton event converts it to a background for like sign dilepton analysis. In this section we describe a data-driven method, developed by myself [72, 102], to measure the probability of charge mis-measurement, which is denoted in the rest of this thesis by P_{cmm}. We

will study the effect of different cuts on this error and try to find the best cuts to reduce wrong-charge electrons and keep the right ones. We will try to use P_{cmm} to estimate number of backgrounds due to this mis-identification fully based on the data.

Measurement of P_{cmm} : Using the $Z \to ll$ events is basically a common method to measure how well the leptons are reconstructed by CMS. Detection of the leptons from Z decay is easier because their invariant mass reproduces the Z-mass peak. So we can find a reliable set of the leptons by applying a cut on the invariant mass of leptons and then try to measure how well a lepton can pass the selection cuts in our analysis. Similar to astronomy, the Z events are usually referred as "*standard candles*" in the experimental high energy physics. In these methods which are called "tag-and-probe", first a lepton is *tagged* with the most reliable cuts. Then if another lepton with looser criteria could make a Z boson, it could be used to *probe* the efficiency of the reconstruction, identification and isolation and even the probability for any kind of error. This way we could measure the efficiencies using the real data and without any need of Monte-Carlo generation and simulation. These methods have a lot of details which are out of the context of this thesis. CMS has used this method to report the performance of the muon [75] and electron [103] in 2010 data.

The Z events can be used to measure the probability of electron charge mis-identification, but in a slightly different way. In a Z event we do not know the charge of which electron is reconstructed correctly to *tag* it. The only thing we know is that the Z events should have two opposite sign leptons. We can use this fact and if an event with two same-sign leptons under the Z-mass peak is found we can assume that the charge of one of them is measured incorrectly. At first we need to make the event selection so tight that only Z events pass it. This selection should also be independent of the charge of the electrons to be able to select same sign events too. In each same-sign event, the charge of one of the leptons is mis-identified. Dividing it by the total number of electrons, the probability of the charge mis-identification can be defined as :

$$P_{cmm} = \frac{N_{ss}}{2 \times (N_{ss} + N_{os})} \tag{40}$$

Where N_{ss} and N_{os} are the number of same-sign and opposite sign Z events. The factor of 2 in the denominator is needed because each Z event has two leptons and so the total number of leptons would be twice the total number of events. N_{ss} events in the

numerator is the number of wrong charge leptons.

The first question is that "how precise is this equation?" and "which approximation is applied in order to obtain this equality?". The only assumption behind this equation is that we assumed the charge of all of the electrons in the opposite sign events are reconstructed correctly. We know that if the charge of both of the electrons in a Z event is mis-identified, the result would be an opposite sign event, but the probability of charge mis-identification is small and the probability of double charge mis-identification for electrons which is P^2_{cmm} is really negligible :

$$P^2_{cmm} \approx 0$$

The only approximation to obtain this equation seems reasonable according to what is expected from the CMS detector.

The shape of P_{cmm} vs. different variables Using the method described in the previous part, only the total error on the charge mis-identification can be measured. This value is also applicable only for the electrons which have a similar kinematics to the electrons of the Z decay. To understand on which variable it depends and for which electrons the charge mis-identification is more probable we need to measure it vs. different variables, like η and p_T.

Extending the method to be able to measure the partial P_{cmm} is not as straight forward as the definition of it. The problem is that from two electrons in a same-sign Z event, we don't know the charge of which one is wrongly measured. So from a same-sign event, we don't know the information (like p_T and η) of which leg can be used in measuring the partial P_{cmm}. In the tag-and-probe like methods, in each pair, there is a lepton which is tagged and we decided to assume it is a correct lepton and the efficiency that we want to measure is the probability of the probe to pass our cuts. So in the tag-and-probe method, the properties like η and p_T of the probe can be used to find the efficiencies and probabilities. But the situation for the charge mis-identification measurement is not that clear.

Each attempt to separate the leptons according to their properties biases the measurement of P_{cmm}. For example if we take the hardest lepton in a same-sign Z event as the mis-measured one, the probability of charge mis-identification for higher values of p_T will be larger.

This problem could be solved by reducing it to a simpler problem : instead of knowing the wrong charge leptons, the p_T or η distribution of them is enough to measure the P_{cmm} vs. these variables. Finding these distributions is possible using the Z events. To justify the formula of (40), we assumed that all of the charge of the leptons in the opposite-sign Z events are reconstructed correctly. So the distribution of all of the leptons in the opposite-sign events can be assumed as the distribution of the leptons for which the charge is correctly assigned.

Half of the leptons in the same-sign Z events have also the right charge. But the charge of the rest is assigned wrongly and the distribution of these is interesting for our study. The distribution of all of the leptons in same-sign events can also be obtained and is the sum of the distribution of well and wrong charge leptons. Thus subtracting the distribution of electrons in opposite sign events with an appropriate weight from the distribution of all electrons in same-sign events results into the distribution of mis-identified electrons.

The final formula with the right weights and coefficients is :

$$N^i_{bad} = N^i_{SS} - \frac{N_{SS}}{2 \times N_{OS}} \times N^i_{OS}$$

Where the i index, represents the i'th bin of the distribution. N^i_{bad} is the number of electrons with mis-identified charge in the i'th bin of the distribution. N^i_{SS} and N^i_{OS} are the contents of the i'th bin of the histogram of all leptons in same-sign and opposite-sign events respectively. N_{SS} and N_{OS} represent the total number of same-sign and opposite-sign events. The η distribution of all of the electrons in same-sign events are compared with the behavior of the electrons in opposite sign events in Figure 48 on page 99.

To derive this formula let's represent the unit normalized distribution of all good electrons with G and for the wrong electrons with F. F and G are the distributions that we don't know explicitly and want to find. The normalized distributions of all leptons in the same-sign and opposite-sign events are know (we take the p_T distribution as an example, it could be η or any other variable) :

$$SS(p_T) = \frac{SS^0(p_T)}{2 \times N_{SS}}$$

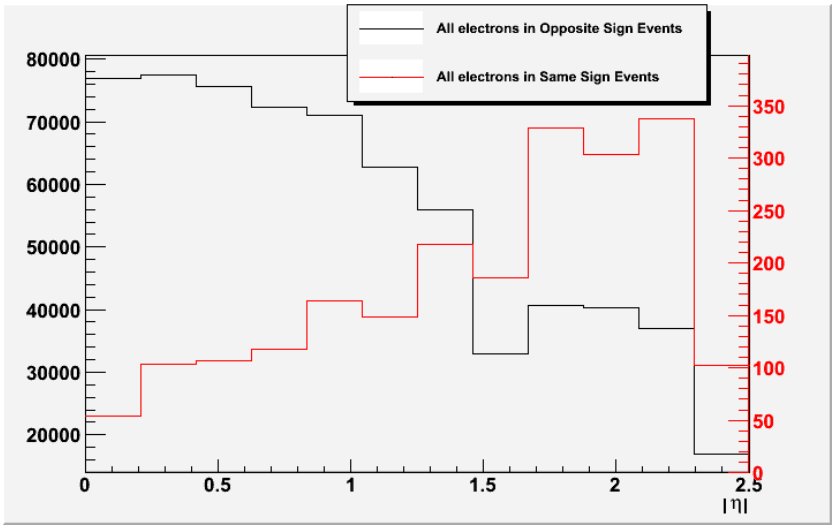


Figure 48: Comparison of the η distribution of the electrons in opposite-sign events (in black) and same-sign events (in red). As each same-sign event contains a mis-identified electrons, so the red curve contains information about the electrons which their charge is mis-measured. It is obvious that the probability of charge mis-identification in the endcap is larger.

$$OS(p_T) = \frac{OS^0(p_T)}{2 \times N_{OS}}$$

Where $SS^0(OS^0)$ is the distribution of all electrons in same-sign (opposite-sign) events. The normalized distributions (OS and SS) are obtained by dividing to the number of electrons in each distribution. We know that the leptons in opposite-sign events are all reconstructed correctly, so we can write :

$$G(p_t) = OS(p_t)$$

On the other hand, we know that the distribution of the electrons in the same-sing events is made up of good and wrong electrons :

$$\frac{1}{2}(F(p_t) + G(p_t)) = SS(p_t)$$

Simplifying these equations, we can easily find that :

$$N_{SS} \times F(p_t) = SS^0(P_t) - \frac{N_{SS}}{2 \times N_{OS}} \times OS^0(p_t)$$

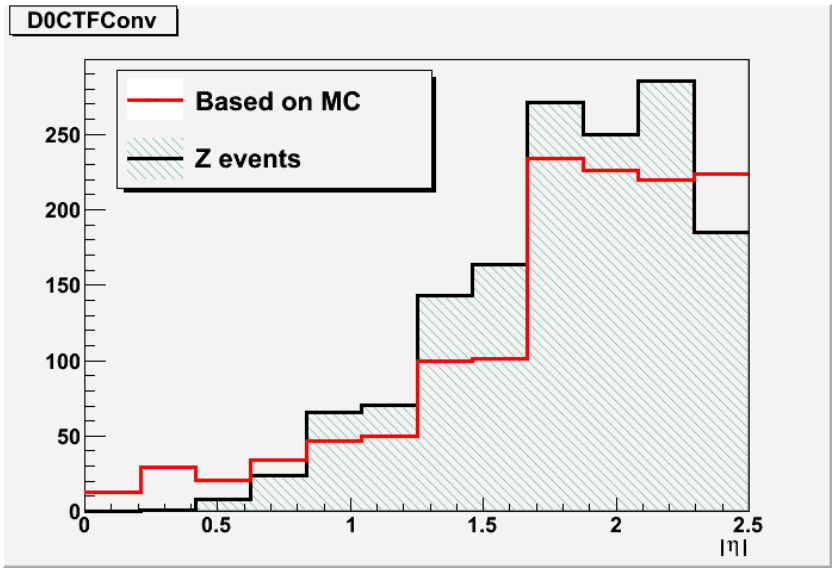


Figure 49: The η distribution of the mis-identified electrons. The curve in black shows the output of the method described in this chapter. The red plot has been obtained using the MC information in the same events. The shapes are in a good agreement.

The left hand side of this equation, is the distribution of the mis-identified electrons (F) weighted by the total number of these electrons in our selection.

After finding the distribution of wrong charge electrons, the next step is to calculate the distribution of all electrons by summing the distribution of all electrons in both same-sign and opposite-sign events. Dividing them together, we can measure the probability of charge mis-identification vs. different variables.

To verify this method, the distribution of the mis-identified electrons in a simulated Z sample by matching the reconstructed electrons with their closest generated electron and comparing their charges have been obtained. Comparing it with the prediction of our method for this distribution is shown in Figure 49 on page 100.

Z selection Now we need to select the $Z \to ee$ events without applying any cut on the charge of the leptons. The Electron dataset is used for this selection. The cuts on the electrons for selecting Z events should be similar to the electron selection of our main analysis. It has two advantages :

1. The probability of charge mis-identification may vary with the identification and isolation of the electron. If the same electron selection is used for the electrons in

our main study is used for Z selection, we can be sure that the resulting P_{cmm} is applicable for our analysis.

2. The selection we have used in our analysis is tighter than the electroweak group uses for Z study [95]. Using a looser set of cuts, more Z events can be selected. But the backgrounds are then important and should be considered carefully. The tighter electron selection that we have used helps us in rejecting a lot of QCD backgrounds.

The first cut for the Z selection is to select all of the events with two electrons and then require their invariant mass to be in a window around the Z-mass. But the probability of charge mis-identification may depend on the momentum of the electrons. The charge and the p_T of an electron is obtained using its bending in the magnetic field of the detector. The radius of the bending of the track of the harder electrons is larger than for softer electrons. Hence for the harder electrons, charge mis-measurement is more likely than for the softer ones.

As the track momentum and charge are both obtained from its curvature, they are very correlated and any error in the measurement of one could mean that the other one is also not well reconstructed. So the uncertainty of the track momenta of the wrong charge electrons in same-sign Z events is expected to be larger than that for other electrons. We can confirm it by drawing the invariant mass of the same-sign events and comparing it to the same distribution for opposite sign events. For calculating invariant mass only the momentum of the tracks of the electrons is used. This comparison in simulation is shown in Figure 50 on page 102. It is seen that the distribution is wider for same-sign events.

As discussed in the electron reconstruction part (4.1), there is another possibility to measure the electron energy instead of the curvature of the track. The momentum of the electron is a combination of the track momentum and the energy of the supercluster. It was also mentioned that for softer electrons, the track momentum is more reliable while for the electrons with higher p_T the energy of the supercluster is more precise. To be independent of the tracker system and not bias the event selection, to select the Z events, the supercluster energy of the electrons is used to calculate the invariant mass. The invariant mass is requested to be between 76 and 106 GeV ($m_Z \pm 15GeV$).

The cuts on the electrons are tight enough to ignore the QCD evens. But the W events need more investigation. In a $W + jets$ events, if the W decays to electron

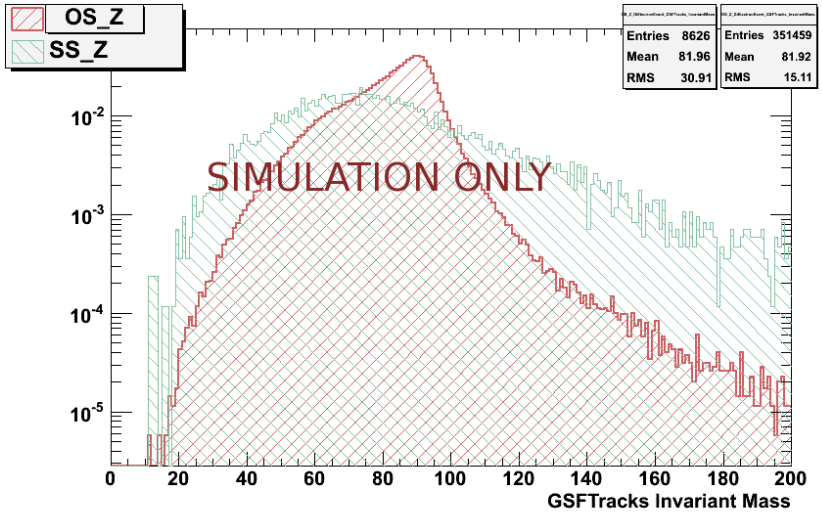


Figure 50: Invariant mass of the same-sing and opposite sign Z events using only the GSF track momentum of the electrons. The distribution for the same-sing events is wider and it confirms that the momentum of the wrong charge electrons is also measured with larger uncertainty.

and neutrino, the electron can easily pass the identification and isolation cuts of our Z selection. It is possible that one of the jets also fakes an electron and passes all of the cuts and fakes a dilepton event. Although the invariant mass of these electrons is not necessarily around Z mass, as the cross section of W production is almost 10 times higher than Z, there could be some $W \to e\nu$ events in the Z events that we have selected.

Applying the Z selection cuts on the simulated W sample and normalizing the result to $35pb^{-1}$, shows that about 12 events can pass the cuts. As the charge of the fake electron could be positive or negative independent of the charge of the real lepton, so half of these events are same-sign and the rest are opposite-sign. 6 same-sign events could affect the P_{cmm}. There are two different variables to separate W events : MET and Transverse mass (M_T). We know that the ν in W events produce real MET while the MET of Z events is not real. Transverse mass of the W events has a peak around the mass of W boson and it could help us to discriminate between W and Z events.

For the current luminosity ($35pb^{-1}$) and to reject only 12 events, we decided to apply only a cut on $MET < 25GeV$. With the remaining events we can measure the probability of charge mis-identification. Applying an extra cut on M_T discards more Z

Table 9: Number of Z events in 2010 data of CMS, with the selection cuts described in the text, MET is requested to be less than $25GeV$. But to make sure that no cut on the M_T is needed, the results with different cuts are reported. Number of events are grouped in same-sign and opposite-sign events. The total probability of charge mis-identification and its statistical error is reported.

M_T	P_{cmm}	N_{SS}	N_{OS}	$\sigma^{statistical}_{P_{cmm}}$
–	0.11%	17	7442	24.24%
45	0.11%	13	6165	27.72%
40	0.10%	12	5744	28.85%
35	0.11%	11	5069	30.13%
30	0.12%	10	4322	31.60%
25	0.10%	7	3522	37.78%

events without any effect on the value of P_{cmm}. To make sure that the value of P_{cmm} is independent of the cut on M_T, the number of remaining events and the value of P_{cmm} for different values of the M_T cut is shown in Table 9 on page 103.

The statistical errors on the probabilities in Table 9 on page 103 are calculated using the binomial confidence interval formula :

$$\delta\hat{\varepsilon} = \sqrt{\frac{\hat{\varepsilon}(1-\hat{\varepsilon})}{N}}$$

Where ε is the value of the probability and N represent the total number of events [104].

It is seen in Table 9 on page 103 that increasing the cut on M_T doesn't affect the value of P_{cmm} significantly, but the statistical error is increased as Z events are discarded.

The P_{cmm} is very low (about 1 per mil 1‰) as expected. There are only 17 mis-identified electrons with which any further investigation is difficult. We can use the described method to find how this probability varies with the transverse momentum of the electron. The result is shown in Figure 51 on page 104. Because of low statistics of the Z events, the error bars are too large and so a firm conclusion from the plot is not possible.

To simplify, we can calculate the probability of charge mis-identification for encap and barrel separately. The same method is used and the results are compared to MC. The result is shown in Figure 52 on page 105. The comparison shows that the method is working well in data. The difference of P_{cmm} in barrel and endcap is larger than

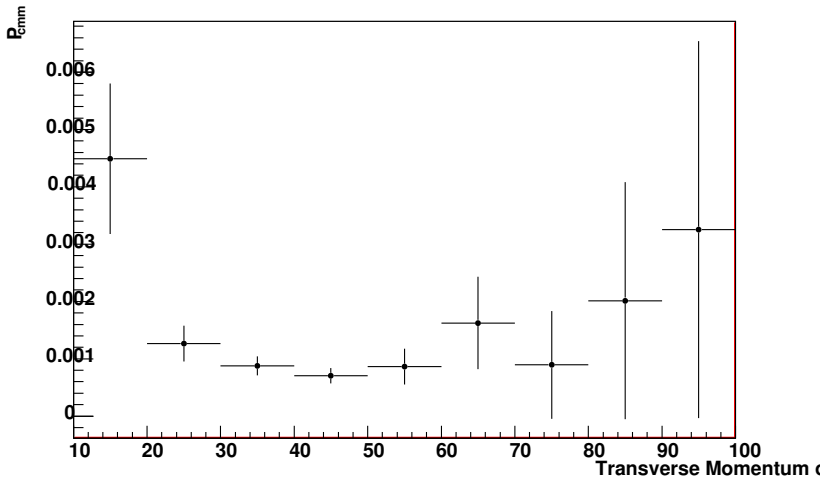


Figure 51: The probability of mis-identification of the electron charge vs. p_T of the electron. The method which is described in the text has been used to obtain this plot.

the error bars and it is obvious that the charge of the electrons in the endcap is more probable to be mis-identified.

Why the electron charge is mis-identified The ability of measuring P_{cmm} vs. different variables helps us to understand why the charge of an electron could be mis-identified.

The effect of the p_T of the electron on the charge mis-identification is understood, the higher p_T of the electron the smaller curvature of its track and so the larger probability for charge mis-identification. But as was shown in the Figure 51 on page 104, the effect of p_T is not so important that it can be taken as the main reason.

The dependence of P_{cmm} on η which is shown in Figure 52 on page 105 was observed in the simulations and seems more important than the momentum of the electron.

Simulation studies show that the main element in mis-reconstructing the charge of an electron is bremsstrahlung. When a high energy photon is emitted by an electron in the very first layers of the tracker, the photon could convert to a pair of electron and positron. The energies of the electron and positron are not necessarily equal. It is possible that the produced positron carries most of the photon energy. So most of the energy of the first electron is now carried by a positron and the track of the positron is reconstructed instead of the main track of the electron which is missing now. A

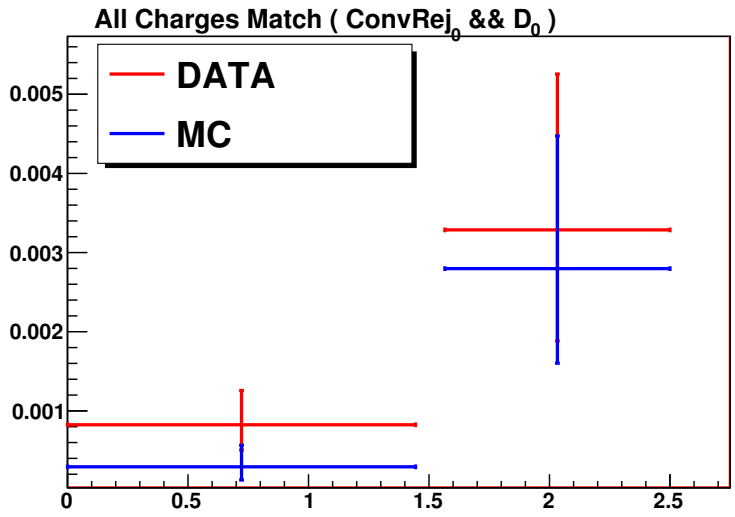


Figure 52: The probability of mis-identification of the electron charge in barrel and endcap. The results of the data-driven method has been compared with MC.

schematic image of this circumstance is shown in Figure 53 on page 106.

The bremsstrahlung happens only in the matter and so its rate depends mostly on the amount of the material that it traverses. The material budget of the CMS tracker vs. η is shown in 22. In the endcaps more material exists. So the probability of bremsstrahlung and hence the probability of charge mis-identification is larger there. It confirms the result that we have found in the previous section.

Methods to reduce the error The probability of charge mis-identification is very low, 0.001 which is reasonable for our purposes. Reaching this small value is possible thanks to the good reconstruction of the electron and also the tight cuts that are applied for electron selection. In this part we want to study the effect of each cut in order to reduce the value of P_{cmm}.

Among all the cuts, one is directly related to the electron charge : requesting all of the three sources of electron charge to agree with each other. As said earlier, there are three methods to measure the charge of an electron : the charge of GSF and CTF track and the $\Delta\phi$ between the innermost position of the track and the supercluster. To assign a CTF track to an electron, we should find the nearest track that has the maximum number of common hits with its GSF track. These variables are calculated at the reconstruction

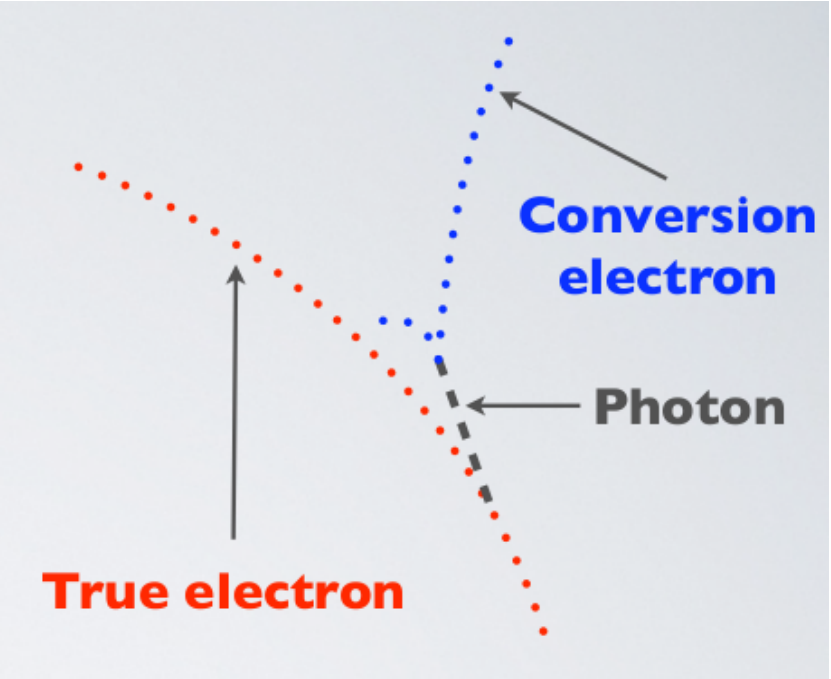


Figure 53: The main known reason for electron charge mis-identification is conversion of an emitted photon in the very first layers of the tracker. If the product with opposite charge carries most of the prompt electron energy, the mis-identification is probable.

level since $CMSSW\_3\_3$ and before that only GSF charge was available. Afterwards the default value of the electron charge has changed to the common value between at least two of these variables. Since the value of the charge can only be positive or negative, finding the majority charge does not discard any electron and each electron will have a proper charge.

Here and in our analysis, we use the electrons with all their charges equal. Such a cut may remove some of the electrons. We expect that the cut only discards mis-identified electrons, but some of the good electrons may also be removed. To study the effect of this cut, we start from GSF charge. All of the cuts for object and event selection are as before except the cut on the charge of the electron. 7920 Z events have been found in 2010 data which is about 6% more than the number of Z events that we reported in the first row of Table 9 on page 103. More Z events are selected, because the selection is looser. If we separate these events into same-sign and opposite sign groups according to the GSF charge of the electrons, we can measure the P_{cmm} for the GSF charge : 169 same-sign events which means $P_{cmm} = 1.07\%$. The relative statistical error is 7.5%. It shows that the GSF charge is ten times less accurate.

Without rejecting any of these events, we can separate them according their majority charge. The number of same-sign events drops to 74 which is equivalent to $P_{cmm} = 0.47\%$ which is less than half of the error when GSF charge is only used.

The results of the tightest cut has been already reported. If we ask all of the charges to be equal, by killing 6% of the good electrons, the P_{cmm} reduces to 0.11% which is small enough for the purpose of our analysis.

Another cut in the electron selection which is expected to be effective to reduce the wrong charge electrons, is "photon conversion rejection". The conversion rejection cuts are primarily applied on electrons to discard the electrons and positrons which are products of photon conversion. These cuts look for another product of the conversion in the track collection. An electron is rejected if the first layers of the tracker are not hit by the electron. As the main reason of charge mis-identification is the conversion of the emitted photons along the path of the electron, the same signatures should be able to discriminate them. We can study the effect of these cuts on the value of P_{cmm} by measuring it before and after applying the cuts.

Using the majority charge and without any cut for conversion rejection, 9808 Z events are found among which 233 same sign events exist and so $P_{cmm} = 1.19\%$. Applying the geometrical conversion rejection reduces this value to $P_{cmm} = 0.85\%$ by discarding 14%

of good electrons. Adding the cut on the number of missing inner hits rejects 4% of the good electrons and reduces the probability of charge mis-identification to 0.47%.

The effect of these cuts is checked for the electrons with three sources of their charges equal, but it is marginal. Without applying any of those conversion rejection cuts $P_{cmm} = 0.29\%$ which is almost three times larger than what was reported in Table 9 on page 103 (P_{cmm} is 0.11% there). These cuts reject 16% of the good electrons.

We also used only GSF and CTF tracks to reconstruct the charge of electrons : electrons are accepted only if the CTF and GSF charges are equal. The results of this step should be compared with the results when only GSF track was used. Using only GSF track, 7920 Z events pass the cuts and $P_{cmm} = 1.07\%$. Requesting the CTF charge and GSF charge to be equal, rejects only 2% of good electrons and reduces the probability of charge mis-identification to 0.31%.

Considering all of the above studies we can say that applying all of the cuts on electrons results to the smallest value for P_{cmm} and is our best choice. The number of good electrons which are discarded after applying these cuts does not increase much by using looser cuts.

Estimating the background due to charge mis-identification The next step which is the main motivation for this study, is the estimation of the backgrounds due to electron charge mis-identification. In fact we want to estimate how many of the 4 observed same sign dilepton events in 2010 data (1 ee and 3 $e\mu$'s) are due to this kind of error.

The low value of P_{cmm} does not mean that this background is small. Because in addition to the value of P_{cmm}, it depends on the number of opposite sign dilepton events which the charge of one of their leptons mis-measured. Hence we need to know how many opposite sign dilepton events with the same selection that applied for same-sign selection exist.

The results of opposite sign selection for different selection regions are shown in Table 10 on page 109. The estimated number of backgrounds from charge flip can be written as

$$N_{SS}^{bkg} \approx P_{cmm} \times (2 \times N_{OS})$$

The factor of two is needed because the charge of each of two leptons in an opposite sign event can be mis-identified. The results are shown in Table 10 on page 109.

Table 10: Number of SS and OS events after applying all of the cuts. The last column includes the estimated number of backgrounds due to charge flip. The results are shown for each selection separately. The errors on the estimation is the propagation of the statistical error of the P_{cmm}

–	$ee\,high\,p_T$	$e\mu\,high\,p_T$	$ee\,low\,p_T$	$e\mu\,low\,p_T$
N_{SS}	0	3	1	0
N_{OS}	15	46	6	5
estimation	0.03 ± 0.01	0.05 ± 0.01	0.01 ± 0.00	0.01 ± 0.00

The estimated value of this kind of background is only a few percent of the number of observed events. In fact the data should be hundred times more than the 2010 data in order to make sure that at least a few events due to charge flip are created.

MC studies show that the main source of this kind of background will be mainly the full leptonic decays of $t\bar{t}$ events, because their topology is similar to the topology of the SUSY events with enough jets and MET. Hence some of them pass the selection cuts that we apply to select SUSY events.

The effect of bining on the estimation In the previous part and to estimate the number of backgrounds due to charge flip, the total probability of charge mis-identification was used. Although the estimation is very small, using partial P_{cmm} vs. different variables may change it.

Two questions should be answered here, which variable works better and what is the effect of the number of bins. For this part of the study, as the 2010 data is not enough, simulated events are used. Z sample to measure the value of P_{cmm} and $t\bar{t}$ sample to validate the method.

The idea it to measure the probability of charge flip from the Z sample vs. η and p_T using the method which is described above. So in the $t\bar{t}$ data sample, there are many opposite-sign events. The probability that the charge of an electron is mis-identified depends on its η or p_T. Estimation is not as simple as total P_{cmm} case. We can estimate the number of same sign events by summing all of these probabilities. Just to validate the method, the number of same-sign ee events among $t\bar{t}$ events is compared to the estimation with using only the total value of P_{cmm} and as it is shown in Figure 54 on page 110 the method prediction is compatible with the number of observed same-sign $t\bar{t}$ events.

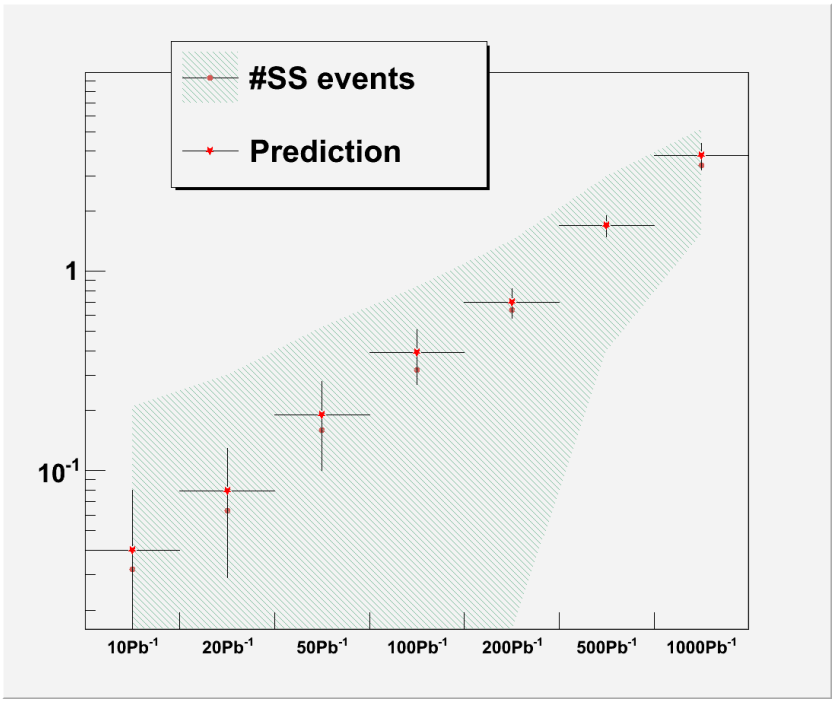


Figure 54: Number of same-sign events among $t\bar{t}$ events compared with our estimation in different integrated luminosities. For estimation the total value of P_{cmm} is used. The statistical errors are all obtained by running pseudo experiments on simulated data.

We can see use the P_{cmm} vs. η and p_T with different binnings and see how the prediction evolves. The estimated numbers in $10pb^{-1}$ are compared in histograms of Figure ?? on page 111. The result is that the estimation does not change a lot and the error bars get larger with increasing the number of bins. Using more bins, the number of Z events in each bin becomes smaller and so the statistical error on the P_{cmm} in each bin becomes larger. The statistical error on P_{cmm} propagates to the error of the estimation and so the more bins we use, the larger error on the estimation.

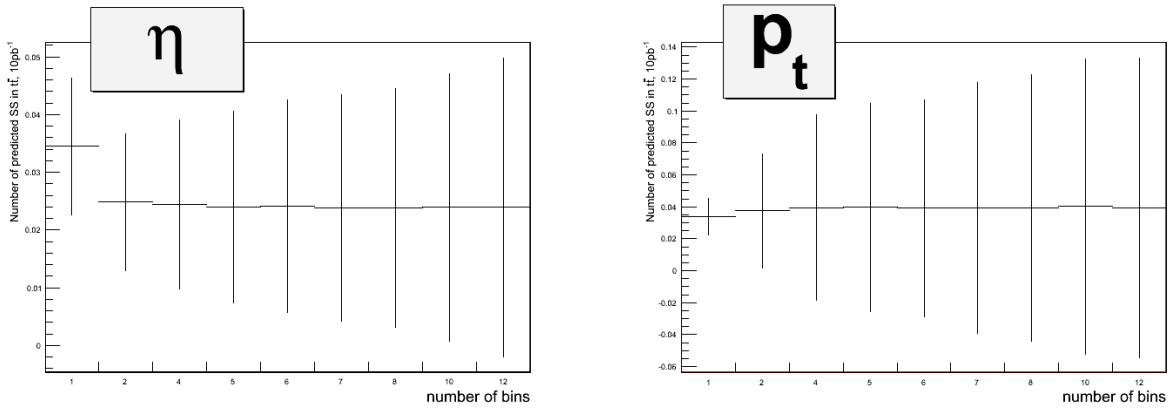


Figure 55: The effect of using different variables and different binnings in the estimation of same-sign events in the $t\bar{t}$ events. The x-axis shows different number of bins. It is seen that the central value of estimation does not change so much, but the error bars increase with increasing the number of bins.

5.3.3 Data-driven estimation of backgrounds containing fake leptons: QCD, W+jet and $t\bar{t}$

Introduction In this part we want to measure the number of backgrounds due to fake leptons in our analysis. We consider as *fake* any lepton that does not originate from a prompt decay such as from W or Z. We can categorize the fake leptons according to their origin :

- Leptons which are produced due to the decay of heavy flavor mesons like B and D mesons. The hadronization of the quarks lead to the production of these mesons.
- The $\pi^{\pm}$ or $K^{\pm}$ in the jets could fake an electron when their energy is completely deposited in the ECAL. Even in some cases muons from K decay can penetrate into the muon system and fake muons.

- In some cases, supercluster of narrow jets becomes so similar to the supercluster of electrons that if a track matches it, the whole jet fakes an electron.

- Another important source for fake electrons is the conversion of photon to electrons. In this process two real leptons are created but they are considered as fake for our purposes.

Prompt leptons are the leptons characterizing the signal, in particular same-sign leptons from chargino decays in SUSY, but can come also from leptonic Z and W boson decays in SM background processes.

Although each of the different kinds of fake leptons have their characteristics, they should be distinguished from prompt leptons by making use of their isolation (ISO) and identification (ID) properties. For example the leptons which are created inside a jet are not as isolated as prompt electrons. The conversion rejection cuts are also used to separate the electrons which are produced from a photon decay.

As none of these selection cuts are perfect, applying all of the lepton selection cuts does not remove all non-prompt leptons. So if in an event with one prompt lepton like the decay of W boson into electron or muon and semi leptonic decays of $t\bar{t}$ events if such a fake lepton is reconstructed a dilepton event is formed. In QCD events no prompt lepton exists, instead there are lots of jets that can fake leptons and it is possible to be reconstructed as a dilepton event too. If by chance the charge of the leptons in these fake events are the same, the event may be counted as a same-sign dilepton background to our analysis.

We want to use a method based on real data to estimate the number of backgrounds due to fake leptons.

Methodology The "fake-ratio" method is a common data-driven method to estimate the contribution of fake lepton backgrounds. The estimation of the fake lepton contribution is based on the generalized fake ratio method, see [105] for a detailed description. Two sets of selection cuts are defined: a *loose* set with rather loose requirements for lepton ID and ISO; and a *tight* set with more stringent requirements for lepton ID and ISO. All other selection criteria (like jet multiplicity, MET, H_T ...) are the same between the loose and the tight selections. The ratio of fake leptons passing the tight criteria over fake leptons passing the loose criteria is called the "fake ratio", f. The similar tight-to-loose ratio for prompt leptons is called the "prompt ratio" p.

The fake ratio f can be determined from abundantly produced QCD events as a function of the lepton kinematics (p_T and η). It is assumed that, once the lepton kinematics is taken into account, the same "universal" fake ratios can be applied to the other backgrounds to extract the signal.

The prompt ratio p can be measured in Z + jets events by using a tag and probe method.

To start with, let us consider the simplest case of single leptons. This is relevant for the $W + jets$ and for the semi-leptonic $t\bar{t}$ channels, contaminated by fake leptons mainly from QCD. The same is true for SUSY single lepton channels.

Suppose that the total number, N_l, of leptons passing the loose criteria is made of N_p prompt and N_f fake leptons. The numbers N_p and N_f are not directly measurable. However, they are related to the numbers of events where no lepton, N_{t0}, or 1 lepton, N_{t1}, pass the tight criteria by

$$\begin{aligned} N_l &= N_p + N_f = N_{t0} + N_{t1} \\ N_{t0} &= (1-p)N_p + (1-f)N_f \\ N_{t1} &= pN_p + fN_f \end{aligned} \tag{41}$$

In these expressions, f is the "fake ratio" and p the "prompt ratio". The ratios f and p depend on the lepton kinematics (p_T and η). The factors can be interpreted as averages over the lepton spectra, specifically over the fake leptons for f and over the prompt leptons for p. Equivalently, they can be used as event by event weights over the fake and prompt leptons. But, we do not know which leptons are fake and which are prompt. What can be measured, instead, is N_{t1} and N_{t0}.

These relations are easily inverted to obtain N_p and N_f, from which the number of events with a prompt lepton and its contamination at the level of the loose selection are derived:

$$\begin{aligned} N_p &= \frac{1}{p-f}\left[(1-f)N_{t1} - fN_{t0}\right] \\ N_f &= \frac{1}{p-f}\left[pN_{t0} - (1-p)N_{t1}\right] \end{aligned} \tag{42}$$

Then, the number of selected prompt (signal) events is given by $N_{signal} = N_p^{pass} = pN_p$ and the number of fakes passing the tight cuts by $N_{contam} = N_f^{pass} = fN_f$.

To extend it to estimate the dilepton events with one or two fake leptons within the signal selection, we define the following quantities:
N_l is the number of events where both leptons pass the loose selection (as mentioned earlier: all other selection cuts are the same between the loose and the tight selection). N_{t2} is the number of events where both leptons, after having passed the loose selection, also pass the tight selection. Similarly, N_{t1} is the number of events where one of the leptons passes the tight selection and one fails. N_{t0} is then the number of events where both leptons fail the tight criteria (still, while having passed the loose ones). These measurable quantities can be related to the "true" number of events containing two prompt, one prompt/one fake, or two fake leptons. Therefore, we define the number of events with both prompt leptons passing the loose criteria as N_{pp}, the number of events with one lepton prompt and one fake as N_{fp} and the number of events where both leptons are fake within the loose criteria as N_{ff}. We then can write

$$\begin{aligned} N_l &= N_{pp} + N_{fp} + N_{ff} = N_{t2} + N_{t1} + N_{t0} \\ N_{t0} &= (1-p)^2 N_{pp} + (1-p)(1-f) N_{fp} + (1-f)^2 N_{ff} \\ N_{t1} &= 2p(1-p) N_{pp} + [f(1-p) + p(1-f)] N_{fp} + 2f(1-f) N_{ff} \\ N_{t2} &= p^2 N_{pp} + pf N_{fp} + f^2 N_{ff} \end{aligned} \tag{43}$$

These equations assume that the prompt and the fake ratios for different leptons are independent of each other. The factors p and $(1-p)$ are weighting (or are averaged over) the distribution of prompt leptons and f and $(1-f)$ are weighing (or are averaged over) the distributions of fake leptons.

After inverting this set of equations, one obtains for the number of events with two prompt leptons:

$$N_{pp} = \frac{1}{(p-f)^2} \left[(1-f)^2 N_{t2} - f(1-f) N_{t1} + f^2 N_{t0} \right] \tag{44}$$

with the number of signal events (i.e. within the tight selection cuts) being given by $N_{signal} = p^2 N_{pp}$. Similarly, one can derive for the number of events with 1 prompt and

1 fake lepton and the number of events with both fake leptons:

$$\begin{aligned} N_{fp} &= \frac{1}{(p-f)^2}\left[-2fpN_{t0} + \left[f(1-p)+p(1-f)\right]N_{t1} - 2(1-p)(1-f)N_{t2}\right] \\ N_{ff} &= \frac{1}{(p-f)^2}\left[p^2N_{t0} - p(1-p)N_{t1} + (1-p)^2N_{t2}\right] \end{aligned} \tag{45}$$

Failing leptons are weighted by f or p and passing leptons by $(1-f)$ or $(1-p)$. As we do not know for N_{t0} which lepton was prompt and which one fake, we weight alternatively lepton 1 with f and 2 with p and lepton 1 with p and 2 with f, adding the weights (hence the factor 2). The same is done with $(1-f)$ and $(1-p)$ for N_{t2}. For N_{t1}, the failing lepton gets alternatively a weight f and p and the passing lepton simultaneously $(1-p)$ and $(1-f)$.

The corresponding backgrounds remaining in the tight selection cuts are then respectively pfN_{fp} and f^2N_{ff}. The first one corresponds to the dominant $t\bar{t}$ background, where the prompt lepton comes from a top decay and the fake lepton from the leptonic b decay of the other top. It might also include a contribution from $W + jets$. The second expression then mainly gives a handle on the remaining QCD contribution.

Fake ratio measurement for muons For muons, the difference between tight and loose lies only in the value of the isolation cut. The tight cuts are defined as described above. For the high-p_T selection we cut on the hybrid isolation smaller than 0.1, for the low-p_T selection we cut on the simple relative isolation smaller than 0.15. In both cases we use a cut value of 1.0 of the respective isolation definitions as the definition of "loose". The ratio is then defined as the number of events, passing the event selection cuts for the following two cases, containing one tight muon to the number of events containing one loose muon. The numerator is always a subset of the denominator, hence the ratio is always ≤ 1.

For the fake ratio we use event selection cuts designed to suppress signal contributions:

- Exactly one loose muon (more muons are vetoed to suppress contribution from Z+jet events)
- Transverse mass between muon and pfMET < 20 GeV , to discard $W \to e\mu$ events.
- Missing transverse energy (pfMET) < 20 GeV

- At least two selected jets (pfJet, $p_T > 30$ GeV)

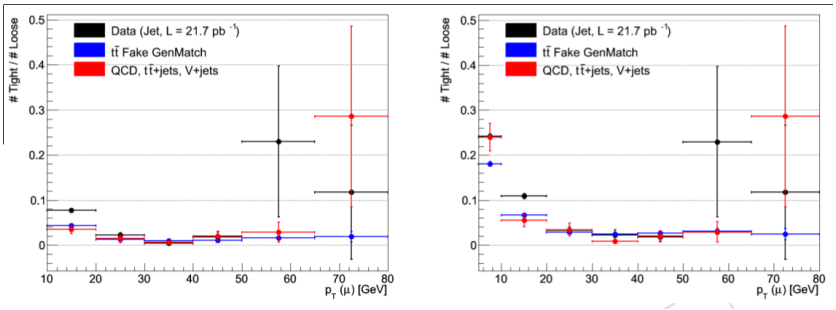


Figure 56: Fake ratio for muons. Left: high p_T selection, right: low p_T selection.

For the prompt ratio we use a selection that is favoring muons from Z decays:

- Exactly two loose muons with opposite charge
- At least two selected jets (pfJet, $p_T > 30$ GeV)
- Invariant mass between the two muons within 15 GeV of the Z boson mass
- Missing transverse energy (pfMET) < 20 GeV
- The ratio is then measured for the harder of the two muons

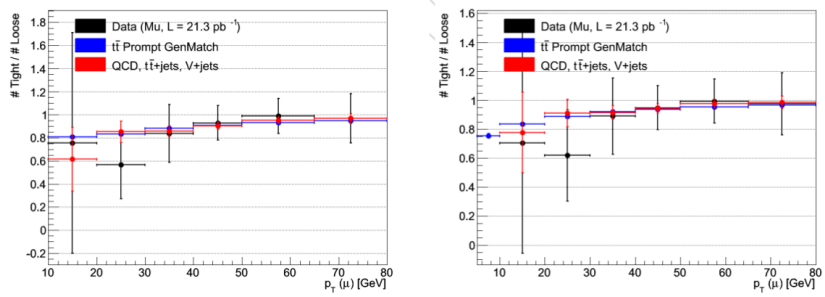


Figure 57: Prompt ratio for muons. Left: high p_T selection, right: low p_T selection.

Fake ratio measurement for electrons The main sources of non-prompt electrons in our analysis are electromagnetic jets matched to a high p_T track (usually originating

from light-flavor quarks) and real non-isolated electrons coming from decays of heavy-flavor (HF) mesons. The fake electrons originating from jets have typically quite different profile of electromagnetic shower than prompt electrons, while the fake electrons coming from HF decays can be distinguished from real electrons by their isolation. For this reason, in order to obtain a reasonably good control on the rate of both of these fake electrons, we have defined a "loose" electron by relaxing both the isolation cut and electron identification (El.ID) cuts. In the two studied selections we define "loose" electrons as those which pass El.ID and conversion rejection cuts which correspond to the so called "90% efficiency working point" (WP90). The cuts of this working point are listed in Table 11 on page 117[96].In addition, in case of the high-p_T selection we require their hybrid isolation to be smaller than 1.0 (0.6 in EE), and for the low-pt selection we require their simple relative isolation to be smaller than 1.0 (0.6 in EE). In both studied selections we require tight electron to pass El.ID and conversion rejection cuts which correspond to the "80% efficiency working point" (WP80). Furthermore, for the high-p_T selection, we cut on the hybrid isolation smaller than 0.1, while for the low-p_T selection we cut on the simple relative isolation smaller than 0.15.

Table 11: The electron identification cuts for $WP90\%$.

	$\frac{H}{E}$	$\Delta\eta_{in}$	$\Delta\Phi_{in}$	$\sigma_{i\eta i\eta}$
Barrel	0.12	0.007	0.8	0.01
Endcaps	0.05	0.009	0.7	0.03

The ratio in a given selection region is then defined as an efficiency for "loose" electrons in that selection region to be selected as tight electrons. Since numerator in this ratio is always a subset of the denominator, the ratio is an efficiency-like ratio that is always ≤ 1.

In order to measure/estimate the electron fake ratio we use the following event selection cuts (designed to mimick the environment in the signal selection region and to suppress contribution of real electrons coming from Z and W decays):

- Exactly one loose electron (more electrons are vetoed in order to suppress contribution from Z+jet events)
- Transverse mass between electron and pfMET $<$ 20 GeV (in order to suppress contribution from W+jet events)

- Missing transverse energy (pfMET) < 20 GeV (in order to suppress contribution from W+jet events)
- At least two selected jets (pfJet, $p_T > 30$ GeV)
- The same HT cut as in the signal selection region ($H_T > 30$ GeV in case of the high-p_T selection and $H_T > 300$ GeV in case of the low-p_T selection)

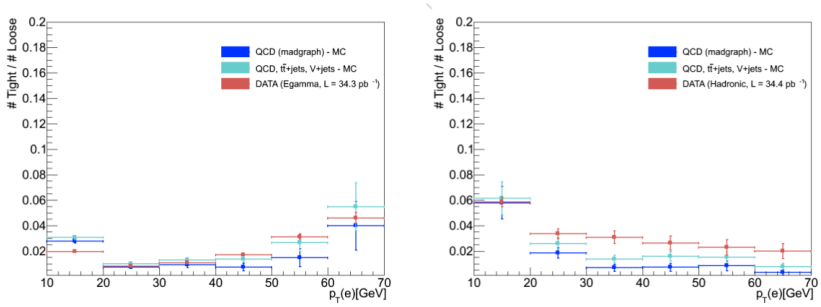


Figure 58: Fake ratio for electrons in the case of the high-p_T (left) and low-p_T (right) selection.

For the electron prompt ratio, as in the case of muon prompt ratios, we use event selection that is favoring Z+jet events:

- Exactly two loose electrons with opposite charge
- Invariant mass between the two electrons within 15 GeV of the Z boson mass
- Missing transverse energy (pfMET) < 20 GeV
- At least two selected jets (pfJet, $p_T > 30$ GeV)
- The ratio is then measured for the harder of the two electrons

Fake rate estimates In case of muons, the fake ratio is measured on the Jet/JetMET dataset, requiring hadronic triggers, using the signal suppressing event selection specified above. The prompt ratio is measured on the Muon/EG/Electron data sets, requiring leptonic triggers, using the Z decay event selection from above. The ratios are no longer binned in pT, but integrated over the whole acceptance, and applied to integrated numbers. The expected event yields in data in 35 pb^{-1} do not allow for a differential

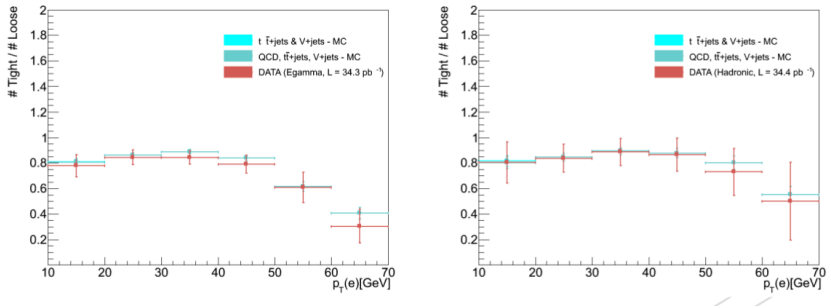


Figure 59: Prompt ratio for electrons in the case of the high-p_T (left) and low-p_T (right) selection.

prediction. Both high-p_T and low-p_T event selections / trigger requirements have been applied for the final signal selection. The resulting predictions are summarized in Tables 12 and 13.

Table 12: Predictions for single- and double-fake lepton backgrounds for the high p_T selection, corresponding to an integrated luminosity of 35 pb^{-1}.

Background	High p_T selection		
	$\mu\mu$	ee	$e\mu$
Fake-fake	0.0458 ± 0.017(stat) ± 0.0231(syst)	0.0355 ± 0.0116(stat) ± 0.0004(syst)	0.29 ± 0.0365(stat) ± 0.124(syst)
Prompt-fake	0.606 ± 0.237(stat) ± 0.129(syst)	0.425 ± 0.172(stat) ± 0.0016(syst)	0.861 ± 0.338(stat) ± 0.249(syst)
Total bkgd.	0.652 ± 0.237(stat) ± 0.131(syst)	0.46 ± 0.172(stat) ± 0.0016(syst)	1.15 ± 0.34(stat) ± 0.279(syst)

In case of the low p_T selection, the electron fake ratio is measured on the Jet/JetMET dataset requiring the same hadronic triggers which are used for the signal selection as well as selection cuts for the suppression of real electrons specified above. Similarly, in case of the high p_T selection, the electron fake ratio is measured on the Electron/EG dataset requiring the same electron triggers as those used in the signal selection region and event selection for real electron suppression specified above. The main motivation for this approach comes form the fact that definitions of several used electron HLT triggers contain tight requirements on electron identification (or even isolation) which bias the effective electron fake rate in the signal selection region.

Table 13: Predictions for single- and double-fake lepton backgrounds for the low p_T selection, corresponding to an integrated luminosity of 35 pb^{-1}.

Background	Low p_T selection		
	$\mu\mu$	ee	$e\mu$
Fake-fake	0.227 ± 0.115 (stat) ± 0.15(syst)	0.0037 ± 0.0064(stat) ± 0.0004(syst)	0.0426 ± 0.0332(stat) ± 0.0366(syst)
Prompt-fake	0.0259 ± 0.414(stat) ± 0.142(syst)	0.114 ± 0.12(stat) ± 0.0056(syst)	0.793 ± 0.439(stat) ± 0.244(syst)
Total bkgd.	0.253 ± 0.429(stat) ± 0.206(syst)	0.118 ± 0.12(stat) ± 0.0056(syst)	0.833 ± 0.441(stat) ± 0.247(syst)

5.4 Validation

To be able to come to a firm conclusion, we need to validate all of the results that have been obtained and the methods that have been used so far. We used the generated and simulated MC samples of standard model processes. The list of all of the samples is shown in 14. The LM0 sample as a sample SUSY point is also used.

The results of the high p_T and low p_T same-sign dilepton selection on the samples are shown in Table 15 on page 122, separated by the decay channel.

The simulation results confirm that the expected number of standard model processes is compatible with the observed data in each selection and they have also been estimated correctly using the data-driven methods which are described in previous parts.

5.5 SUSY Exclusion

No evidence of new physics has been observed in 2010 data of CMS in the same-sign dilepton channel. We can set limits on new physics with the observed data.

In the published results by CMS [106], the mentioned results in ee, $\mu\mu$ and $e\mu$ channel are combined with the τ lepton channels and the limits are set with the combined result. For background estimation, other methods are also used in CMS to confirm all aspects of the analysis. Knowing the number of observed events and backgrounds and all the errors, we can set the 95%CL upper limit on the number of observed events using a Bayesian method[107].

To exclude a particular model, we can generate and simulate data with that model, apply all of the cuts and compare the number of remaining events with the upper

Table 14: List of CMSSW_3_6 MC samples used in this study, and their cross sections.

Process	Sample name	σ (pb)
QCD (Pythia)	/QCD_Pt15/Spring10-START3X_V26_S09-v1 (with $\widehat{p_T} < 30$ GeV)	8.76e+08
	/QCD_Pt30/Spring10-START3X_V26_S09-v1 (with $\widehat{p_T} < 80$ GeV)	6.04e+07
	/QCD_Pt80/Spring10-START3X_V26_S09-v1 (with $\widehat{p_T} < 170$ GeV)	9.24e+05
	/QCD_Pt170/Spring10-START3X_V26_S09-v1	2.55e+04
QCD (Madgraph)	/QCD_Pt-50To100_7TeV-madgraph/Spring10-START3X_V26-v1	3e+07
	/QCD_Pt100to250-madgraph/Spring10-START3X_V26_S09-v2	7e+06
	/QCD_Pt250to500-madgraph/Spring10-START3X_V26_S09-v1	1.71e+05
	/QCD_Pt500to1000-madgraph/Spring10-START3X_V26_S09-v1	5.2e+03
	/QCD_Pt1000toInf-madgraph/Spring10-START3X_V26_S09-v1	8.3e+01
TTBar	/TTbarJets-madgraph/Spring10-START3X_V26_S09-v1	157.5
WJets	/WJets-madgraph/Spring10-START3X_V26_S09-v1	3.13e+04
ZJets	/ZJets-madgraph/Spring10-START3X_V26_S09-v1	3.05e+03
Astar	/AstarJets-madgraph/Spring10-START3X_V26_S09-v1	310
VVJets	/VVJets-madgraph/Spring10-START3X_V26_S09-v1	4.8
SSW+W+	/SSWWSPSPos_7_Digi/volper-SPSPos_reco_7TeV	0.188
SSW-W-	/SSWWSPSNeg_7_Digi/volper-SSWW_SPSNeg_Reco_7TeV	0.064
SSWW DPS	/SSWWDPS_7_Digi/volper-SSWWDPS_Reco_7TeV	0.203
LM0	/LM0/Spring10-START3X_V26_S09-v1	38.93

limit. This job has been done in CMS for a grid of fast simulated SUSY samples in the mSUGRA plane and the exclusion line in the $m_0 - m_{1/2}$ is shown in Figure 60 on page 122.

CMS also provided a generic method that every new physics theory could use to exclude some part of its parameter phase space. In order to do that, the efficiency of signal selection is needed. The main factor in the efficiency is the efficiency of lepton selection. In the published result of CMS, the best fit on the efficiency of lepton selection is given as a function of p_T of the lepton, so any other theory can use it to see if it is excluded. To validate this efficiency function, the mSUGRA exclusion curve is recomputed without simulating the CMS experiment. The results are also shown in Figure 60 on page 122.

Table 15: List of yields for the signal and various background processes, as well as the data yield, corresponding to an integrated luminosity of 35 pb^{-1}.

MC/data sample	High p_T selection			Low p_T selection		
	$\mu\mu$	ee	$e\mu$	$\mu\mu$	ee	$e\mu$
$t\bar{t}$	0.23	0.15	0.41	0.13	0.03	0.18
W+jets	0.00	0.22	0.33	0.000	0.00	0.00
Z+jets	0.00	0.00	0.10	0.000	0.00	0.00
VV+jets	0.04	0.04	0.16	0.01	0.002	0.01
QCD	0.00	0.00	0.00	0.000	0.00	0.00
$qq \to qqW^+W^+$	0.03	0.02	0.054	0.02	0.009	0.027
$qq \to qqW^-W^-$	0.01	0.007	0.016	0.005	0.004	0.006
double parton $W^\pm W^\pm$	0.001	0.001	0.003	0.00	0.00	0.00
Total background	0.31	0.43	1.1	0.16	0.041	0.22
LM0 (LO)	2.0	1.3	3.3	2.2	1.0	3.1
Data	0	0	3	0	1	0

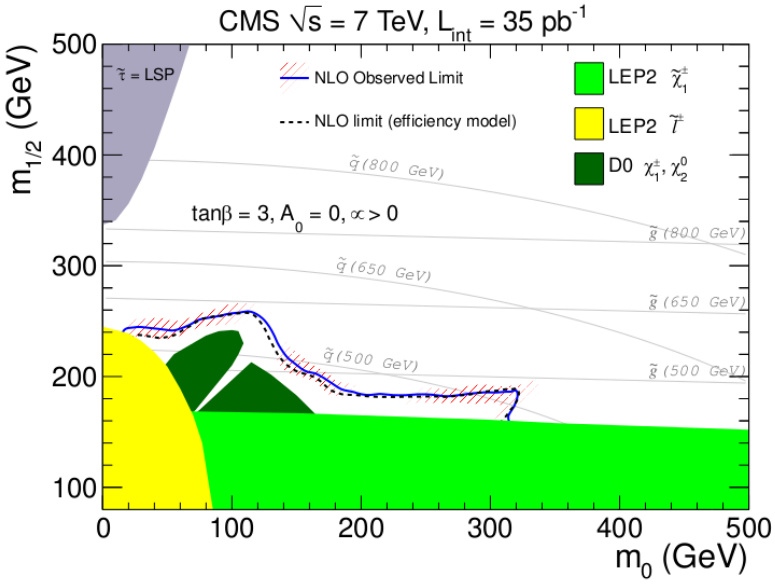


Figure 60: The exclusion line in the $m_0 - m_{1/2}$ plane. The solid blue line is obtained by simulating the SUSY in different points of this plane. The dashed black line is also set using the proposed lepton efficiency function.

A CMSSW Module to select high purity tracks

The CMSSW config file which selects the high quality tracks, can be found under the RecoTracker/FinalTrackSelectors/python/selectHighPurity_cfi.py file in CMSSW. A copy is shown here :

```
import FWCore.ParameterSet.Config as cms

selectHighPurity = cms.EDProducer("AnalyticalTrackSelector",
    src = cms.InputTag("generalTracks"),
    keepAllTracks = cms.bool(False), ## if set to true tracks failing th
    beamspot = cms.InputTag("offlineBeamSpot"),

    # vertex selection
    useVertices = cms.bool(True),
    vertices = cms.InputTag("pixelVertices"),
    vtxNumber = cms.int32(-1),
    vertexCut = cms.string('ndof>=2&!isFake'),

    #untracked bool copyTrajectories = true // when doing retracking bef
    copyTrajectories = cms.untracked.bool(False),
    copyExtras = cms.untracked.bool(True), ## set to false on AOD
    qualityBit = cms.string('highPurity'), ## set to '' or comment out i

    # parameters for adapted optimal cuts on chi2 and primary vertex con
    chi2n_par = cms.double(0.7),
    res_par = cms.vdouble(0.003, 0.001),
    d0_par1 = cms.vdouble(0.3, 4.0),
    dz_par1 = cms.vdouble(0.35, 4.0),
    d0_par2 = cms.vdouble(0.4, 4.0),
    dz_par2 = cms.vdouble(0.4, 4.0),
    # Boolean indicating if adapted primary vertex compatibility cuts ar
    applyAdaptedPVCuts = cms.bool(True),

    # Impact parameter absolute cuts.
```

```
    max_d0 = cms.double(100.),
    max_z0 = cms.double(100.),
    nSigmaZ = cms.double(4.),

    # Cuts on numbers of layers with hits/3D hits/lost hits.
    minNumberLayers = cms.uint32(3),
    minNumber3DLayers = cms.uint32(3),
    maxNumberLostLayers = cms.uint32(2),

    # Absolute cuts in case of no PV. If yes, please define also max_d0N
    applyAbsCutsIfNoPV = cms.bool(False)

)
```

B ETH Physics quality checking

For handling the info form and showing the image of each data set according to its status, a python script has been developed.

B.1 Installation

1. Apache httpd and Python should be installed on the server
2. Mod_python (> 3.2) is needed to integrate python scripts into httpd.
   - in SLC it's very simple : *yum install mod_python*
3. Mod_python should be configured
   - The simplest way to do that is by adding the following lines to */etc/httpd/conf.d/python.conf*

     ```
     AddHandler modpython .py \\PythonHandler modpython.publisher
     \\PythonDebug On
     ```

   - For more information you can read the online documentation of Mod_python

4. These files should be copied to the ETHZPhysQC home directory :

   - *FormHandler.py* which is the main file
   - *info_files* directory which contains *TheForm.html* which is the template of the form.
   - *img* directory which contains *blue.gif*, *red.gif*, *white.gif*, *green.gif* and *yellow.gif*.

5. The *info_files* directory should be writable by *apache* user, and all of its files should be deletable by *physqc* members, a solution is like this :

   - *apache* should be the member of *physqc*
   - The directory permission should be like this : 2770

     ```
     chmod 2770 infofiles
     ```

     By this settings, the owner group of all of the files created under this directory will be *physqc*. So all of its members can delete them.

B.2 Functionalities

B.2.1 Loading the form

This method load the info form for a given data set

The only parameter it needs, is *dsname* and should pass as Query String. Its value should be the name of the directory in which the data set histograms are.

To load the info form, you can easily call
http://wwweth.cern.ch/ETHPhysQC/FormHandler.py/Load?dsname=_dsname_

It loads the form in 4 steps :

1. It opens the template form file from info_files/TheForm.html

```
htmlfile = os.path.join(directory,'info_files/TheForm.html')
formFile = open(htmlfile , 'r')

rettem = Template(formFile.read())
```

```
formFile.close()
```

2. It reads the *dsname/info.txt*. This file is supposed to be produced during the production of histograms. Its firs line contains the *date and time* of the production and its third line contains the number of events which are analyzed.

```
timeOfTheRun = 'N/A'
nEvents = '-1'

infofile = os.path.join(directory,dsname + '/info.txt')
try:
    info = open(infofile , 'r')
    timeOfTheRun = info.readline()
    info.readline()
    nEvents = info.readline()
    info.close()
except:
    nEvents = 'N/A'
```

3. It reads the info_files/dsname. If it doesn't exist, it means that no information is available for this data set. But if it is there, the first line shows the status of the data set. As it's mentioned earlier, status can be either GOOD or BAD of DUBIOUS. The second line and third line are the last modification time and the name of the last user who commented. The rest of the file is the comments for this data set.

```
infofile = os.path.join(directory,'info_files/' + dsname )
try:
        info = open(infofile , 'r')
        status = info.readline()

        GOODCHECKED = ''
        DUBIOUSCHECKED = ''
        BADCHECKED = ''
```

```
            if  str(status)[0:-1]==str('GOOD'):
                GOODCHECKED = 'checked=\"checked\"'
            elif  str(status)[0:-1]=='BAD':
                BADCHECKED = 'checked=\"checked\"'
            elif  str(status)[0:-1]=='DUBIOUS':
                DUBIOUSCHECKED = 'checked=\"checked\"'
            modificatedAt = info.readline()
            name = info.readline()
            comment = info.readline()
            line = info.readline()
            while line:
                comment += line
                line = info.readline()

            info.close()

            allVals = dict(GOODCHECKED=GOODCHECKED,
                           BADCHECKED=BADCHECKED,
                           DUBIOUSCHECKED=DUBIOUSCHECKED,
                           comment=comment,
                           name=name,DSNAME=dsname,
                           TIME=timeOfTheRun,NEvents=nEvents,
                           LastUpdate=modificatedAt)
            return str(rettem.safe_substitute( allVals ))
    except IOError:
            allVals = dict(GOODCHECKED='',
                           BADCHECKED='',DUBIOUSCHECKED='',
                           comment='',name='',DSNAME=dsname,
                           TIME=timeOfTheRun,
                           NEvents=nEvents,LastUpdate='Never')

            return str(rettem.safe_substitute( allVals ))
```

4. Finally, it updates the Template Form with the values which have been read from

the files and sends the form to the client.

B.2.2 Saving the information

After filling the information form and pressing the send button, it's submitted to the **Save** method. In addition to the dsName which is sent to the Save method by query string, all of the form values are sent.

The Save method checks if the Status and Name fields are filled. If one of them is not set, It shows an error. In other case, it creates the info_files/dsName file (if it already exists, overwrites it). Then writes the values in the file in the given order. Finally it redirects to the Load method for the working dsName.

```
def Save(req, dsName,  name, comment ):
        status = req.form.getfirst('status')
        # make sure the user provided all the parameters
        if not (dsName and status and name):
                return "A required parameter is missing, \
                                please go back and correct the error"
        directory = os.path.dirname(req.filename)
        outfile = os.path.join(directory,'info_files/' + dsName)
        f=open(outfile, 'w')
        f.write(status + "\n")
        now = datetime.datetime.now()
        f.write( now.strftime("%d %b %Y, %H:%M:%S") + '\n' )
        f.write(name+ "\n")
        f.write(comment)
        f.close()
        util.redirect(req, 'Load?dsname=' + dsName)
    return outfile
```

B.2.3 Display the dataset icon

For displaying the status icon of each data set, an specific method has been developed. It can simply be called via
http://wwweth.cern.ch/ETHPhysQC/FormHandler.py/GetImage?dsname=_dsName_
Where dsName is the name of the data set you want to see its status icon.

The only thing that this method does is to open the info_files/dsname and read the first line and detect the status of the data set. Then according to the value, redirects the page to image.

```
from mod_python import apache
def GetImage(req, dsname):
        req.content_type = "image/gif"
        req.send_http_header()

        directory = os.path.dirname(req.filename)
        infofile = os.path.join(directory,'info_files/' + dsname )
        try:
                info = open(infofile , 'r')
                status = str(info.readline())[0:-1]
                info.close()
                if status=='GOOD':
                        util.redirect(req, '../img/green.gif')
                elif status=='BAD':
                        util.redirect(req, '../img/red.gif')
                elif status=='DUBIOUS':
                        util.redirect(req, '../img/yellow.gif')
                else:
                        util.redirect(req, '../img/blue.gif')
        except IOError:
                util.redirect(req, '../img/white.gif')
    return apache.OK
```

References

[1] Emmy Noether and M. A. Tavel. "Invariant Variation Problems". In: *Gott.Nachr. Statist.Phys.1:186-207,1971* Gott.Nachr.1918:235-257,1918; Statist.Phys.1:186-207,1971 (Gott.Nachr.1918:235-257,1918; Transp.Theory Statist.Phys.1:186-207,1971), Gott.Nachr.1918:235–257,1918;Transp.Theory Statist.Phys.1:186–207,1971. eprint: physics/0503066(Cited on p. 1).

[2] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.508(Cited on p. 1).

[3] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.321(Cited on p. 1).

[4] T. W. B. Kibble. "Symmetry Breaking in Non-Abelian Gauge Theories". In: *Phys. Rev.* 155 (5 1967), pp. 1554–1561. DOI: 10.1103/PhysRev.155.1554. URL: http://link.aps.org/doi/10.1103/PhysRev.155.1554(Cited on p. 1).

[5] Rabindra N. Mohapatra and Goran Senjanovi ć. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (14 1980), pp. 912–915. DOI: 10.1103/PhysRevLett.44.912. URL: http://link.aps.org/doi/10.1103/PhysRevLett.44.912(Cited on p. 5).

[6] P. J. E. Peebles and Bharat Ratra. "The cosmological constant and dark energy". In: *Rev. Mod. Phys.* 75 (2 2003), pp. 559–606. DOI: 10.1103/RevModPhys.75.559. URL: http://link.aps.org/doi/10.1103/RevModPhys.75.559(Cited on p. 5).

[7] Th. M. Nieuwenhuizen. "Do non-relativistic neutrinos constitute the dark matter?" In: *EPL (Europhysics Letters)* 86.5 (2009), p. 59001. URL: http://stacks.iop.org/0295-5075/86/i=5/a=59001(Cited on p. 5).

[8] Ben Moore et al. "Dark Matter Substructure within Galactic Halos". In: *The Astrophysical Journal Letters* 524.1 (1999), p. L19. URL: http://stacks.iop.org/1538-4357/524/i=1/a=L19(Cited on p. 5).

[9] Oliver Sim BrÃŒning et al. *LHC Design Report v.1 : the LHC Main Ring*. Geneva: CERN, 2004. URL: http://cdsweb.cern.ch/record/782076(Cited on pp. 5, 29).

[10] Stephen P. Martin. "A Supersymmetry Primer". In: (2011). eprint: hep-ph/9709356(Cited on pp. 5, 9, 13, 16, 19).

[11] Tao Han et al. "Phenomenology of the little Higgs model". In: *Phys. Rev. D* 67 (9 2003), p. 095004. DOI: 10.1103/PhysRevD.67.095004. URL: http://link.aps.org/doi/10.1103/PhysRevD.67.095004(Cited on p. 6).

[12] J. Wess and B. Zumino. "Supergauge transformations in four dimensions". In: *Nuclear Physics B* 70.1 (1974), pp. 39 –50. ISSN: 0550-3213. DOI: 10.1016/0550-3213(74)90355-1. URL: http://www.sciencedirect.com/science/article/pii/0550321374903551(Cited on p. 8).

[13] Rabindra N. Mohapatra and Goran Senjanovi ć. "Neutrino Mass and Spontaneous Parity Nonconservation". In: *Phys. Rev. Lett.* 44 (14 1980), pp. 912–915. DOI: 10.1103/PhysRevLett.44.912. URL: http://link.aps.org/doi/10.1103/PhysRevLett.44.912(Cited on p. 11).

[14] Irene Niessen. "Supersymmetric Phenomenology in the mSUGRA Parameter Space". In: (Sept. 2008). eprint: 0809.1748(Cited on pp. 15, 17).

[15] L. Girardello and M.T. Grisaru. "Soft breaking of supersymmetry". In: *Nuclear Physics B* 194.1 (1982), pp. 65 –76. ISSN: 0550-3213. DOI: 10.1016/0550-3213(82)90512-0. URL: http://www.sciencedirect.com/science/article/pii/0550321382905120(Cited on p. 15).

[16] S. Heinemeyer, W. Hollik, and G. Weiglein. "Precise Prediction for the Mass of the Lightest Higgs Boson in the MSSM". In: *Phys.Lett. B* 440 (1998), pp. 296–304. eprint: hep-ph/9807423(Cited on p. 19).

[17] Charles M. Ankenbrandt et al. "Status of muon collider research and development and future plans". In: *Phys. Rev. ST Accel. Beams* 2.8 (1999), p. 081001. DOI: 10.1103/PhysRevSTAB.2.081001(Cited on p. 22).

[18] S Baird. *Accelerators for pedestrians; rev. version*. Tech. rep. AB-Note-2007-014. CERN-AB-Note-2007-014. PS-OP-Note-95-17-Rev-2. CERN-PS-OP-Note-95-17-Rev-2. Geneva: CERN, 2007(Cited on p. 23).

[19] L Arnaudon et al. "Measurement of the mass of the Z boson and the energy calibration of LEP". In: *Phys. Lett. B* 307.CERN-PPE-93-53. CERN-SL-93-17-AP. FSU-SCRI-94-12. CERN-L3-058 (1993), 187–193. 9 p(Cited on p. 25).

[20] D Bloch et al. *Measurement of the mass and width of the W Boson in e^+e^- Collisions at $\sqrt{s} \simeq$ 192-202 GeV. oai:cds.cern.ch:996523*. Tech. rep. DELPHI-2001-006 MORIO-CONF-448. CERN-DELPHI-2001-006 MORIO-CONF-448. Geneva: CERN, 2001(Cited on p. 25).

[21] J A Wear. "Measurement of the number of light neutrino generations, Z resonance parameters, and absolute luminosity at the ALEPH detector. oai:cds.cern.ch:248609". Presented on 21 Mar 1991. PhD thesis. Geneva: Wisconsin Madison Univ., 1991(Cited on p. 25).

[22] A A Rademakers. "The determination of the number of light neutrino species. oai:cds.cern.ch:217094". Presented on 30 Jan 1991. PhD thesis. Amsterdam: Amsterdam Univ., 1991(Cited on p. 25).

[23] "LEP SUSY Working group". URL: http://lepsusy.web.cern.ch/lepsusy/(Cited on p. 25).

[24] F. Abe et al. "Observation of top quark production in $\bar{p}p$ collisions". In: *Phys.Rev.Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv:hep-ex/9503002 [hep-ex](Cited on p. 26).

[25] S. Abachi et al. "Observation of the top quark". In: *Phys.Rev.Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv:hep-ex/9503003 [hep-ex](Cited on p. 26).

[26] CDF Collaboration. "Invariant Mass Distribution of Jet Pairs Produced in Association with a W Boson in $p\bar{p}$ Collisions at $s = 1.96TeV$". In: *Phys. Rev. Lett.* 106.17 (2011), p. 171801. DOI: 10.1103/PhysRevLett.106.171801(Cited on p. 26).

[27] D0 Collaboration. "Study of the dijet invariant mass distribution in $\boldsymbol{p\bar{p} \to W(\to \ell\nu) + jj}$ final states at $\boldsymbol{\sqrt{s} = 1.96}$ TeV". In: (2011). arXiv:1106.1921. arXiv:1106.1921 [hep-ex](Cited on p. 26).

[28] G. Arnison et al. "Further evidence for charged intermediate vector bosons at the SPS collider". In: *Physics Letters B* 129.3-4 (1983), pp. 273 –282. ISSN: 0370-2693. DOI: DOI:10.1016/0370-2693(83)90860-2. URL: http://www.sciencedirect.com/science/article/pii/0370269383908602(Cited on p. 29).

[29] G. Arnison et al. "Experimental observation of lepton pairs of invariant mass around 95 GeV/c2 at the CERN SPS collider". In: *Physics Letters B* 126.5 (1983), pp. 398 –410. ISSN: 0370-2693. DOI: DOI:10.1016/0370-2693(83)90188-0. URL: http://www.sciencedirect.com/science/article/pii/0370269383901880(Cited on p. 29).

[30] Michael Benedikt et al. *LHC Design Report v.3 : the LHC Injector Chai*. Geneva: CERN, 2004. URL: http://cdsweb.cern.ch/record/823808(Cited on p. 29).

[31] CERN DG. *Summary of the analysis of the 19 September 2008 incident at the LHC*. Tech. rep. Geneva: CERN, 2008. URL: http://cdsweb.cern.ch/record/1135729/(Cited on p. 30).

[32] CERN Press Release. "Follow up of the incident of 19 September 2008 at the LHC". In: (). URL: https://edms.cern.ch/document/973073/1(Cited on p. 30).

[33] G Flugge. "Future Research in High Energy Physics". In: (1994). URL: http://cdsweb.cern.ch/record/401268?ln=en(Cited on p. 31).

[34] CERN LHC Group. "LHC Commissioning with Beam". A web page containing information about 2009 and 2010 LHC runs. URL: http://lhc-commissioning.web.cern.ch/lhc-commissioning/(Cited on p. 32).

[35] James Pinfold et al. *Technical Design Report of the MoEDAL Experiment*. Tech. rep. CERN-LHCC-2009-006. MoEDAL-TDR-001. Geneva: CERN, 2009. URL: http://cdsweb.cern.ch/record/1181486?ln=en(Cited on p. 34).

[36] *CMS Physics Technical Design Report Volume I: Detector Performance and Software*. CERN, 2006(Cited on pp. 35, 37, 44, 45).

[37] Filip Moortgat. "Discovery Potential Of MSSM Higgs Bosons using supersymmetric decay modes with the CMS detector". PhD thesis. UNIVERSITEIT ANTWERPEN, 2004(Cited on p. 40).

[38] V Karimaki. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: http://cdsweb.cern.ch/record/368412?ln=en(Cited on p. 41).

[39] Laura Borrello et al. *Sensor Design for the CMS Silicon Strip Tracker*. Tech. rep. CMS-NOTE-2003-020. Geneva: CERN, 2003. URL: http://cdsweb.cern.ch/record/687861?ln=en(Cited on p. 42).

[40] Serguei Chatrchyan et al. "Commissioning of the CMS Experiment and the Cosmic Run at Four Tesla". In: *JINST* 5 (2010), T03001. DOI: 10.1088/1748-0221/5/03/T03001. arXiv:0911.4845 [physics.ins-det](Cited on pp. 42, 52).

[41] Steven Lowette. *Operation, calibration and performance of the CMS silicon tracker*. Tech. rep. CMS-CR-2010-258. CERN-CMS-CR-2010-258. Geneva: CERN, 2010. URL: http://cdsweb.cern.ch/record/1358641?ln=en(Cited on p. 43).

[42] Vardan Khachatryan et al. "CMS Tracking Performance Results from early LHC Operation". In: *Eur.Phys.J.* C70 (2010), pp. 1165–1192. DOI: 10.1140/epjc/s10052-010-1491-3. arXiv:1007.1988 [physics.ins-det](Cited on p. 43).

[43] *The CMS electromagnetic calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: http://cdsweb.cern.ch/record/349375?ln=en(Cited on p. 44).

[44] CMS Collaboration. "Electromagnetic calorimeter calibration with 7 TeV data". In: (2010). CMS-PAS-EGM-10-003. URL: http://cdsweb.cern.ch/record/1279350?ln=en(Cited on p. 45).

[45] A. Tumasyan et al. Phys. V. Khachatryan A. M. Sirunyan. "Transverse-Momentum and Pseudorapidity Distributions of Charged Hadrons in pp Collisions at $\sqrt{s} = 7$". Rev. Lett. 105 (Jul, 2010) 022002.(Cited on p. 46).

[46] CMS Collaboration. "Electromagnetic calorimeter commissioning and first results with7 TeV data". CMSNOTE 2010 012(Cited on pp. 47, 48).

[47] G Baiatian et al. *Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges*. Tech. rep. CMS-NOTE-2006-138. CERN-CMS-NOTE-2006-138. Geneva: CERN, 2007. URL: http://cdsweb.cern.ch/record/1049915?ln=en(Cited on p. 49).

[48] G Baiatian et al. *Design, Performance, and Calibration of CMS Hadron Endcap Calorimeters*. Tech. rep. CMS-NOTE-2008-010. CERN-CMS-NOTE-2008-010. Geneva: CERN, 2008. URL: http://cdsweb.cern.ch/record/1103003?ln=en(Cited on p. 49).

[49] CMS Collaboration. *HCAL performance from first collisions data*. Tech. rep. CMS DP 2010 025. CMS, 2010. URL: http://cdsweb.cern.ch/record/1364222?ln=en(Cited on pp. 50, 51).

[50] CMS Collaboration. "Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data". In: *Journal of Instrumentation* 5.03 (2010), T03012. URL: http://stacks.iop.org/1748-0221/5/i=03/a=T03012(Cited on p. 50).

[51] CMS Collaboration. *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: http://cdsweb.cern.ch/record/331056?ln=en(Cited on p. 51).

[52] CMS Collaboration. *The CMS muon project: Technical Design Report*. Technical Design Report CMS CERN-LHCC-97-032 ; CMS-TDR-003. Geneva: CERN, 1997. URL: http://cdsweb.cern.ch/record/343814?ln=en(Cited on p. 52).

[53] Serguei Chatrchyan et al. "Fine Synchronization of the CMS Muon Drift-Tube Local Trigger using Cosmic Rays". In: *JINST* 5 (2010), T03004. DOI: 10.1088/1748-0221/5/03/T03004. arXiv:0911.4904 [physics.ins-det](Cited on p. 54).

[54] Serguei Chatrchyan et al. "Performance of the CMS Drift Tube Chambers with Cosmic Rays". In: *JINST* 5 (2010), T03015. DOI: 10.1088/1748-0221/5/03/T03015. arXiv:0911.4855 [physics.ins-det](Cited on p. 54).

[55] Serguei Chatrchyan et al. "Performance of the CMS Drift-Tube Local Trigger with Cosmic Rays". In: *JINST* 5 (2010), T03003. DOI: 10.1088/1748-0221/5/03/T03003. arXiv:0911.4893 [physics.ins-det](Cited on p. 54).

[56] Serguei Chatrchyan et al. "Calibration of the CMS Drift Tube Chambers and Measurement of the Drift Velocity with Cosmic Rays". In: *JINST* 5 (2010), T03016. DOI: 10.1088/1748-0221/5/03/T03016. arXiv:0911.4895 [physics.ins-det](Cite on p. 54).

[57] Serguei Chatrchyan et al. “Performance of the CMS Cathode Strip Chambers with Cosmic Rays”. In: *JINST* 5 (2010), T03018. DOI: 10.1088/1748-0221/5/03/T03018. arXiv:0911.4992 [physics.ins-det](Cited on p. 54).

[58] Serguei Chatrchyan et al. “Performance Study of the CMS Barrel Resistive Plate Chambers with Cosmic Rays”. In: *JINST* 5 (2010), T03017. DOI: 10.1088/1748-0221/5/03/T03017. arXiv:0911.4045 [physics.ins-det](Cited on p. 54).

[59] Manfred Paulini. “CDF – Run II status and prospects”. In: *Nuclear Physics B - Proceedings Supplements* 120 (2003). Proceedings of the 8th International Conference on B-Physics at Hadron Machines, pp. 287 –294. ISSN: 0920-5632. DOI: DOI:10.1016/S0920-5632(03)01917-0. URL: http://www.sciencedirect.com/science/article/pii/S0920563203019170(Cited on p. 54).

[60] Reinhard Schwienhorst. “The D0 Run II Trigger System”. In: *International Journal of Modern Physics A (IJMPA)* 20 (2005), pp. 3796–3798. DOI: 10.1142/S0217751X05027643. URL: http://www.worldscinet.com/ijmpa/20/2016/S0217751X05027643.html(Cited on p. 54).

[61] Sergio Cittolin, Attila RÃ¡cz, and Paris Sphicas. *CMS trigger and data-acquisition project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2002. URL: http://cdsweb.cern.ch/record/578006?ln=en(Cited on p. 55).

[62] Serguei Chatrchyan et al. “Performance of the CMS Level-1 Trigger during Commissioning with Cosmic Ray Muons”. In: *JINST* 5 (2010), T03002. DOI: 10.1088/1748-0221/5/03/T03002. arXiv:0911.5422 [physics.ins-det](Cited on p. 55).

[63] Serguei Chatrchyan et al. “Commissioning of the CMS High-Level Trigger with Cosmic Rays”. In: *JINST* 5 (2010), T03005. DOI: 10.1088/1748-0221/5/03/T03005. arXiv:0911.4889 [physics.ins-det](Cited on p. 55).

[64] Malgorzata Kazana. *CMS trigger and data taking in 2010*. Tech. rep. CMS-CR-2011-051. CERN-CMS-CR-2011-051. Geneva: CERN, 2011. URL: http://cdsweb.cern.ch/record/1358192?ln=en(Cited on p. 55).

[65] I. Antcheva et al. “ROOT – A C++ framework for petabyte data storage, statistical analysis and visualization”. In: *Computer Physics Communications* 180.12 (2009). 40 YEARS OF CPC: A celebratory issue focused on quality software for high performance, grid and novel computing architectures, pp. 2499 –

2512. ISSN: 0010-4655. DOI: DOI:10.1016/j.cpc.2009.08.005. URL: http://www.sciencedirect.com/science/article/pii/S0010465509002550(Cited on p. 58).

[66] W Adam et al. "Track reconstruction in the CMS tracker". In: (2006). CMS NOTE-2006/041(Cited on p. 59).

[67] W Adam et al. "Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC". In: (2005). CMS AN-2005/001(Cited on p. 59).

[68] H. Bethe and W. Heitler. "On the Stopping of Fast Particles and on the Creation of Positive Electrons". In: *Proceedings of the Royal Society of London. Series A* 146.856 (1934), pp. 83–112. DOI: 10.1098/rspa.1934.0140. eprint: http://rspa.royalsocietypublishing.org/content/146/856/83.full.pdf+html. URL: http://rspa.royalsocietypublishing.org/content/146/856/83.short(Cited on p. 60).

[69] Stephanie Baffioni et al. *Electron reconstruction in CMS*. Tech. rep. CMS-NOTE-2006-040. CERN-CMS-NOTE-2006-040. Geneva: CERN, 2006. URL: http://cdsweb.cern.ch/record/934070?ln=en(Cited on p. 62).

[70] David Wardrope. "Electron charge determination using SuperCluster position relative to pixel seed direction". Presented in EGamma POG meeting at CERN. Oct. 2007(Cited on p. 62).

[71] *CMSSW source code for calculating $\sigma_{\eta\eta}$. RecoEcal/EgammaCoreTools/src/ClusterShapeAlgo.cc*(Cited on p. 63).

[72] H. Bakhshiansohi, L. Pape, and Filip Moortgat. "A data driven method to measure electron charge mis-identification rate". In: *1st IPM Meeting On LHC Physics*. 2009(Cited on pp. 63, 95).

[73] N Neumeister. "Muon Reconstruction Software in CMS". In: (2005). URL: http://cdsweb.cern.ch/record/865826(Cited on p. 65).

[74] J. Alcaraz N. Amapane E. Antillon R. Bellan I. Belotelov I. Bloch C. Campagnari T. Cox A. Everett A. Grelli J. Goh V. Halyo A. Hunt E. James P. Kalavase S.C. Kao M. Konecki D. Kovalskyi V. Krutelyov C. Liu P. Martinez D. Miller M. Mulders N. Neumeister D. Pagano J. Pivarski J. Ribnik S. Stoynev P. Traczyk D. Trocino J.R. Vlimant R. Wilkinson G. Abbiendi N. Adam. *Muon Reconstruction in the CMS detector*. Tech. rep. CMS Analysis Note, 2008. URL: http://cms.c

ern.ch/iCMS/jsp/db_notes/showNoteDetails.jsp?noteID=CMS%20AN-2008/097(Cited on p. 65).

[75] "Performance of muon identification in pp collisions at s**0.5 = 7 TeV". In: CMS-PAS-MUO-10-002 (2010). URL: http://cdsweb.cern.ch/record/1279140?ln=en(Cited on pp. 65, 96).

[76] CMS Collaboration. "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET". In: (2009). CMS-PAS-PFT-09-001. URL: http://cdsweb.cern.ch/record/1194487?ln=en(Cited on p. 66).

[77] CMS Collaboration. "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV". In: (2010). CMS-PAS-PFT-10-002. URL: http://cdsweb.cern.ch/record/1279341?ln=en(Cited on pp. 66, 67, 69).

[78] Florian Beaudette. *Performance of the particle flow algorithm in CMS*. Tech. rep. CMS-CR-2010-276. CERN-CMS-CR-2010-276. Geneva: CERN, 2010(Cited on p. 67).

[79] CMS Collaboration. "Jet Energy Corrections determination at 7 TeV". CMS-PAS-JME-10-010. 2010. URL: http://cdsweb.cern.ch/record/1308178?ln=en(Cited on pp. 67, 68).

[80] Howard Baer et al. "Detecting gluinos at hadron supercolliders". In: *Phys. Rev. D* 36.1 (1987), p. 96. DOI: 10.1103/PhysRevD.36.96(Cited on p. 71).

[81] Howard Baer, Xerxes Tata, and Jeffrey Woodside. "Multilepton signals from supersymmetry at hadron supercolliders". In: *Phys. Rev. D* 45.1 (1992), pp. 142–160. DOI: 10.1103/PhysRevD.45.142(Cited on p. 71).

[82] R. Michael Barnett, John F. Gunion, and Howard E. Haber. "Discovering supersymmetry with like-sign dileptons". In: *Physics Letters B* 315.3-4 (1993), pp. 349–354. ISSN: 0370-2693. DOI: DOI:10.1016/0370-2693(93)91623-U. URL: http://www.sciencedirect.com/science/article/pii/037026939391623U(Cited on pp. 71, 74).

[83] W. Beenakker et al. "Squark and gluino hadroproduction". In: (May 2011). eprint: 1105.1110(Cited on pp. 71, 72).

[84] Manoranjan Guchait and D. P. Roy. "Like-sign dilepton signature for gluino production at the CERN LHC including top quark and Higgs boson effects". In: *Phys. Rev. D* 52.1 (1995), pp. 133–141. DOI: 10.1103/PhysRevD.52.133(Cited on p. 74).

[85] Howard Baer et al. "Signals for minimal supergravity at the CERN Large Hadron Collider. II. Multilepton channels". In: *Phys. Rev. D* 53.11 (1996), pp. 6241–6264. DOI: 10.1103/PhysRevD.53.6241(Cited on p. 74).

[86] W. T. Giele and S. Keller. "Determination of W-boson properties at hadron colliders". In: *Phys. Rev. D* 57.7 (1998), pp. 4433–4440. DOI: 10.1103/PhysRevD.57.4433(Cited on p. 75).

[87] C. G. Lester and D. J. Summers. "Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders". In: *Physics Letters B* 463.1 (1999), pp. 99 –103. ISSN: 0370-2693. DOI: DOI:10.1016/S0370-2693(99)00945-4. URL: http://www.sciencedirect.com/science/article/pii/S0370269399009454(Cited on p. 76).

[88] Alan Barr, Christopher Lester, and Phil Stephens. "A variable for measuring masses at hadron colliders when missing energy is expected; m T 2 : the truth behind the glamour". In: *Journal of Physics G: Nuclear and Particle Physics* 29.10 (2003), p. 2343. URL: http://stacks.iop.org/0954-3899/29/i=10/a=304(Cited on p. 76).

[89] Michael Burns et al. "Using subsystem M T 2 for complete mass determinations in decay chains with missing energy at hadron colliders". In: *Journal of High Energy Physics* 2009.03 (2009), p. 143. URL: http://stacks.iop.org/1126-6708/2009/i=03/a=143(Cited on p. 77).

[90] T. Aaltonen et al. "Measurement of the top quark mass in the dilepton channel using $mT2$ at CDF". In: *Phys. Rev. D* 81.3 (2010), p. 031102. DOI: 10.1103/PhysRevD.81.031102(Cited on p. 77).

[91] cms collaboration. "Measurement of the $t\bar{t}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s}$ =7 TeV. oai:cds.cern.ch:1354581". In: (2011)(Cited on p. 77).

[92] Konstantin T Matchev et al. "Precision sparticle spectroscopy in the inclusive same-sign dilepton channel at LHC". In: (2009), p. 4. URL: http://arxiv.org/abs/0909.4300(Cited on p. 77).

[93] R. N. Mohapatra and J. C. Pati. ""Natural" left-right symmetry". In: *Phys. Rev. D* 11.9 (1975), pp. 2558–2561. DOI: 10.1103/PhysRevD.11.2558(Cited on p. 78).

[94] Francisco del Aguila et al. "Heavy Majorana neutrinos in the effective Lagrangian description: Application to hadron colliders". In: *Physics Letters B* 670.4-5 (2009), pp. 399 –402. ISSN: 0370-2693. DOI: DOI:10.1016/j.physletb.2008.11.031. URL: http://www.sciencedirect.com/science/article/pii/S0370269308013981(Cited on p. 78).

[95] Vardan Khachatryan et al. "Measurements of Inclusive W and Z Cross Sections in pp Collisions at sqrt(s)=7 TeV". In: *JHEP* 1101 (2011). * Temporary entry *, p. 080. DOI: 10.1007/JHEP01(2011)080. arXiv:1012.2466 [hep-ex](Cited on pp. 82, 101).

[96] *https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaWorkingPoints*. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaWorkingPoints(Cited on pp. 82, 117).

[97] S. Gessner P. Kalavase D. Kovalskyi V. Krutelyov J. Ribnik D. Barge C. Campagnari. "Study of photon conversion rejection at CMS". In: *CMS Internal Note* CMS AN -2009/159 (2009). URL: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2009/159(Cited on p. 84).

[98] Tom Melia et al. "W^+W^+ plus dijet production in the POWHEGBOX". In: (Feb. 2011). eprint: 1102.4846. URL: http://arxiv.org/abs/1102.4846(Cited on p. 95).

[99] Tom Melia. "W^+W^+ jj at NLO in QCD: an exotic Standard Model signature at the LHC". In: (May 2011). eprint: 1105.5394. URL: http://arxiv.org/abs/1105.5394(Cited on p. 95).

[100] Tom Melia et al. "Next-to-leading order QCD predictions for $W^+W^+ + jj$ production at the LHC". In: *Journal of High Energy Physics* 2010 (12 2010). 10.1007/JHEP12(2010)0 pp. 1–17. DOI: 10.1007/JHEP12(2010)053(Cited on p. 95).

[101] Jonathan R. Gaunt et al. "Same-sign W pair production as a probe of double parton scattering at the LHC". In: (Mar. 2010). eprint: 1003.3953. URL: http://arxiv.org/abs/1003.3953(Cited on p. 95).

[102] A. Calderon P. Lobelle J. Marco P. Martinez T. Rodrigo R. Vilar G. Cappello M. Chiorboli A. Tricomi J. Cuevas J. Fernandez L. LLoret G. Dissertori P. Milenovic F. Moortgat P. Nef L. Pape F. Ronga L. Sala B. Stieger K. Theofilatos D. Treille H. Bakhshian A. Fahim and M. Weber. "Same-Sign di-lepton search for supersymmetry in the first 35 pb-1 of 7 TeV pp collisions". CMS AN-2010/375. 2010. URL: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2010/375(Cited on p. 95).

[103] CMS Collaboration. "Electron reconstruction and identification at sqrt(s) = 7 TeV". In: CMS-PAS-EGM-10-004 (2010). URL: http://cdsweb.cern.ch/record/1299116?ln=en(Cited on p. 96).

[104] Robert D. Cousins, Kathryn E. Hymes, and Jordan Tucker. "Frequentist Evaluation of Intervals Estimated for a Binomial Parameter and for the Ratio of Poisson Means". In: *Nuclear Instruments and Methods in Physics Research A 388–398* Nuclear 388–398 (May 2009), NuclearInstrumentsandMethodsinPhysicsResearchA612(2010) 388–398. eprint: 0905.3831(Cited on p. 103).

[105] A. Calderon P. Lobelle J. Marco P. Martinez T. Rodrigo R. Vilar G. Cappello M. Chiorboli A. Tricomi J. Cuevas J. Fernandez L. LLoret G. Dissertori P. Milenovic F. Moortgat P. Nef L. Pape F. Ronga L. Sala B. Stieger K. Theofilatos D. Treille H. Bakhshian A. Fahim and M. Weber. "Computing the contamination from fakes in leptonic final states". CMS AN-2010/261. 2010. URL: http://cms.cern.ch/iCMS/jsp/db_notes/noteInfo.jsp?cmsnoteid=CMS%20AN-2010/261(Cited on p. 112).

[106] CMS Collaboration. "Search for new physics with same-sign isolated dilepton events with jets and missing transverse energy at the LHC". In: *Journal of High Energy Physics* 2011 (6 2011). 10.1007/JHEP06(2011)077, pp. 1–47. ISSN: 1029-8479. URL: http://dx.doi.org/10.1007/JHEP06(2011)077(Cited on p. 120).

[107] K Nakamura and Particle Data Group. "Review of Particle Physics". In: *Journal of Physics G: Nuclear and Particle Physics* 37.7A (2010), p. 075021. URL: http://stacks.iop.org/0954-3899/37/i=7A/a=075021(Cited on p. 120).

Summary

Operating at $7\,TeV$ center of mass energy in 2010 data taking, the Large Hadron Collider delivered a considerable amount of data, corresponding to an integrated luminosity of $47\,pb^{-1}$, to its experiments. The CMS experiment recorded an integrated luminosity of $\approx 35pb^{-1}$ from the delivered collisions. This amount of data was enough to rediscover all of the aspects of the Standard Model, from weak interactions to QCD, although the Higgs boson is still not observed.

The analyses at the level of reconstructing the physics objects, exploited all capabilities of the CMS tracker and calorimeter. This led to an efficient and very high precision object reconstruction. The detector worked so better than expected that even with the first few inverse pico barns of data the collaboration started to search for new physics in which part the effort was dedicated to the same-sign dilepton signature. This signature in the framework of SUSY was studied in this thesis.

The event selection was performed for two different scenarios. In both scenarios, a high missing transverse energy and hadronic activity (H_T) were required to reject more backgrounds while representing the environment of the expected SUSY signals. These scenarios were optimized according to the available triggers. The first scenario with relatively harder leptons was based on leptonic trigger selections. For the second scenario with softer lepton selection, H_T triggers were used. The total event yield from the full dataset of 2010 operation (35 pb^{-1}) was equal to 4 containing 3 $e\mu$ events by the first and 1 ee event by the second selection. To understand if these events are signals of new physics or are expected from the Standard Model, three possible backgrounds were considered :

- Real same-sign dilepton event from the Standard Model
  There are very rare sources of same-sign dilepton events in the Standard Model. We have shown that for an integrated luminosity of 35 pb^{-1} such contributions are completely negligible.

- Electron charge mis-identification
  In an opposite-sign dilepton event, if the charge of one of the electrons is mis-identified, the event is considered as a background to the same-sign channel. A data driven method to measure the probability of charge mis-identification was developed. The probability was measured by data and found to be around 1 per

mil. The background due to this error was also estimated to be $\approx 0.05(0.02)$ of events for high(low) lepton p_T selection.

- Fake leptons
  A dilepton event is wrongly reconstructed by either events with one prompt lepton and a jet faking a lepton, or a QCD event with two jets faking two leptons. Such events are considered as backgrounds if the (mis-)reconstructed leptons are like-sign. We developed a new method to measure the fake ratio in data and apply it to estimate the background due to the fake. It was found that most of the selected events are from this kind of background.

The results were compared with simulation and a good agreement was observed. So we concluded that there was no evidence for new physics from the same-sign dilepton channel in the data collected by CMS in 2010. Using this information, new bounds on new physics were set.