SEARCH FOR HIGH MASS RESONANCES DECAYING TO TAU PAIRS AT THE CMS EXPERIMENT

By

Carlos Andres Florez Bustos

Dissertation
Submitted to the Faculty of the
Graduate School of Vanderbilt University
in partial fulfillment of the requirements for
the degree of
DOCTOR OF PHILOSOPHY
in
Physics
December, 2011
Nashville, Tennessee

Approved:
Professor Will Johns
Professor Paul Sheldon
Professor Thomas Kephart
Professor Andreas Berlind
Professor Robert Weller

I would like to dedicate this thesis to my beloved wife. This is the first step for a better life together.

ACKNOWLEDGMENTS

I would like to thank the National Science Foundation (NSF) of the United States of America for providing the funding to the high energy physics group at Vanderbilt University that paid for my education and supported me and my family all these years. I have an eternal gratitude to the American people for helping me to make my dreams come true.

I also would like to thank my advisor Professor Will Johns for all his teaching, help, support and useful advice during my Ph.D. I do not have a way to thank him enough for all his help and valuable teaching. I hope to be a good mentor and an example for my future students as he has been for me.

I wish to thank very much all the members of the high energy physics group at Vanderbilt for their help throughout the years. Specially, I would like to thank Professors Sheldon, Webster and Gabella for their valuable explanations that helped me to become a better scientist.

I would like to acknowledge the high energy physics group at Texas A&M University. I thank Professor Alexei Safonov for helping me so much with the Z' analysis and for opening the doors of his institution, hosting me for several months. Also, I thank very much Professor Teruki Kamon for giving me the opportunity to work with him in the SUSY analysis and to have so much fun doing physics. I learned many valuable things from him for my professional life and as a human being.

I would like to give a very special recognition to Dr Alfredo Gurrola for all his help, support and teaching. I thank him so much for guiding my analysis and spending so many hours explaining to me all sorts of physics concepts. He is a great physics mentor and I hope to keep progressing and applying all that I learned from him and also help others as he helped me.

I would like to thank my wife, my parents and my sister for their love and unconditional support.

Finally, I would like to thank my fellow graduate student and friend Michael Clemens for his help all these years. I thank him so much for being like a brother to me and helping me overcome so many difficulties during all this time in graduate school. I will be thankful forever.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

Science is a tool created to answer some of the fundamental questions regarding natural phenomena. Curiosity is the motor that drives hundreds of years of scientific research seeking to explain why we are here and how the Universe works. The study of fundamental particles attempts to answer some of the most profound questions in modern science: how is matter formed? what is dark matter? are there extra dimensions in the universe?

The Large Hadron Collider (LHC) is a proton-proton and heavy ion collider located near Geneva Switzerland. Currently, the LHC is providing a new energy frontier to recreate conditions present in the early Universe in order to understand some of the physics processes that govern the fundamental particles we observe today. The Compact Muon Solenoid (CMS) is a multipurpose detector located at the LHC, designed to reconstruct the particles produced after the collisions of the accelerated particles. The aim of CMS is to find new particles that can reveal signs of new physics and guide us towards an understanding of the underlying natural principles that govern the fundamental particles. One of the main goals of CMS is to see if the Higgs boson, a particle theoretically predicted to be responsible for the mass generation of all particles, exists or not.

Our understanding and experimental evidence about particles has been summarized in a physics model known as the Standard Model (SM) [1]. There are multiple theoretical models that aim to answer some of the still open questions in the SM. Some of these models are being experimentally tested at the LHC. A very interesting topic is the prediction of high mass resonances in different physics models. The discovery of one of these resonances will be direct proof of new physics. In this paper, we present the search for high mass resonances decaying to tau leptons at CMS. We use a Z' particle with standard model couplings as a benchmark in order to develop and tune our analysis. Nevertheless, this is a general search to any high mass

resonance decaying to taus. The decay to taus is favored in several physics models that use different couplings than the standard model [2]. Therefore, taus present an opportunity to discover new physics. If new resonances are found in the other final states such as electrons or muons, it is still important to measure the decay to tau leptons in order to test if lepton universality (the interaction of gauge bosons and leptons is the same for the three known lepton families) is preserved or not [3].

The tau particle is the heaviest lepton and can decay to other leptons and associated neutrinos or to (a jet of) hadrons and a tau neutrino. We present the search for a high mass resonance decaying to a tau pair, where one tau decays leptonically to a tau neutrino, an electron and electron anti-neutrino, and the other tau decays to a tau neutrino and a jet of hadrons. Since taus have different decay modes, we combined our results with other analyses searching in different tau decay modes in order to achieve a better sensitivity and statistical significance. The analyses were carried out by different groups from several Universities [4]. Our combined result exceeds the sensitivity achieved by a previous search carried out by the CDF experiment at the Fermi National Laboratory in Batavia Illinois [5].

CHAPTER II

BASIC CONCEPTS IN PARTICLE PHYSICS

Particle physics is the study of the elementary constituents of matter and their interactions. Several elementary particles have been discovered and some of their interactions have been quantified and qualified, nevertheless there are many open questions to be answered. An elementary particle is defined as a particle not composed of other particles. Three of the four fundamental known interactions are included in the SM: Electromagnetic, Weak and Strong interaction (gravity is not yet included in the SM).

In the SM, matter is composed of twelve elementary particles known as fermions which are divided into two subgroups: leptons and quarks. Fermions are spin half particles that satisfy Fermi-Dirac statistics. The Fermi-Dirac statistics describe the energies of systems of indistinguishable particles that obey the Pauli Exclusion Principle: spin half particles with the same quantum numbers can not occupy the same quantum state. Leptons carry electric charge and quarks carry electric and color charge. As the electric charge is a physical property of charged particles manifested in the electromagnetic interaction, color charge is a physical property of quarks manifested in the strong interaction. According to the theory and experimental evidence, color charge has three possible values which we have named: red, blue and green. The word color in this case is not used literally either, (nobody has seen a quark to tell how it looks!) it describes an extra quantum number that experimental results have shown is needed when quarks interact strongly. The physics processes behind particles with this color "charge" are described with Quantum Chromodynamics (QCD) [1].

Leptons and quarks are organized in families of doublets:

- Quarks: up and down (u, d), charm and strange (c, s) and top and bottom (t, b)
- Leptons: electron and electron neutrino (e, ν_e), muon and muon neutrino (μ, ν_μ) and tau and tau neutrino (τ, ν_τ)

The magnitude of the electric charge of quarks is a fractional value of the electron charge, with values of $(2/3, -1/3)|e|$ for each family respectively. For leptons, the electric charge is $(-1, 0)|e|$ for each family. In the SM the interactions between particles are carried by bosons which are particles with integer spin that satisfy Bose-Einstein statistics. Bose-Einstein statistics describe systems of indistinguishable particles that might occupy the same energy level even if they have the same quantum numbers.

In the SM each particle has an antiparticle. Antiparticles are objects that have the same mass as the particle but opposite electric and/or color charge.

Particles containing two or three quarks form a new group known as Hadrons. These are divided into two subgroups according to their quark content: baryons and mesons. Baryons are particles containing a set of three quarks and mesons are particles that contain a quark and anti-quark pair. For example Protons and Neutrons are baryons, and Pions and Kaons are mesons. The experimental discovery of the $\Delta^{++}(uuu)$ baryon suggested a violation of the Pauli Exclusion Principle. According to the QCD theory only colorless hadrons are permitted and the inclusion of the color charge as a quantum number solved the apparent violation. A colorless particle can be composed of a quark and anti-quark pair (mesons) carrying color and anti-color charge of the same kind respectively (i.e red and anti-red) or three quarks or anti-quarks (baryons) carrying different color or anti-color charge.

II.1 Interactions

As mentioned previously, there are four known fundamental interactions that have been verified experimentally. All particles experience gravitational acceleration. Nevertheless, due to the low mass of the particles that we study gravitational effects are negligible. Gravity is the weakest of all four interactions, but its effects are significant for systems with a large number of atoms, making it the dominant interaction in much of the macroscopic world. Extensions of the SM propose that

the gravitational interaction is mediated by a massless boson of spin two called the graviton which hasn't been found yet. Hence, gravity is an open question in the SM.

Particles that carry electric charge interact electromagnetically. The interaction is carried through the photon (γ) a massless boson that is defined as the minimum unit of electromagnetic radiation.

The weak interaction acts in radioactive processes, *i.e.* beta decays where an electron or positron is emitted. The carriers of the weak interaction happen to be heavy massive bosons of spin one identified as W^+, W^- and Z^0. The charged W is responsible for quark flavor changing, where quarks decay within and across families. The Z^0 participates in scattering processes between fermions and strong interactions that produce a quark and anti-quark pair.

Finally, the strong interaction keeps the quarks bound which results in quarks existing only within particles such as Protons, Neutrons, Pions, Kaons etc. The mediators of the strong interaction are eight bosons, assumed to be massless, called gluons. According to QCD theory, gluons carry color and anti-color charge simultaneously. The color charge for each of the eight gluons is represented as a superposition of color and anti-color charges *i.e*:

$$(r\overline{b} + \overline{r}b)/\sqrt{2} \tag{1}$$

Table 1 summarizes the fundamental properties of the particles that we have described so far.

II.2 Symmetry Groups and Symmetry Breaking in the Standard Model

According to experimental evidence all particles and the three fundamental interactions considered in the SM (electromagnetic, weak and strong) appear to have three internal symmetries described through the unitary groups: $U(1)$, $SU(2)$, and $SU(3)$. The $U(1)$ group, is associated with the invariance of the fields under a phase transformation, for example $\phi \Rightarrow \phi^{'} = \exp^{i\theta}\phi$. Similarly, all particles have a second internal symmetry which make them invariant under transformations described by

Table 1
Fundamental properties of quarks, leptons and intermediate bosons [6].

Particle	$Q(\|e\|)$	Spin	Mass
u	2/3	1/2	1.5 to 3.3 MeV
d	-1/3	1/2	3.5 to 6.0 MeV
c	2/3	1/2	$1.27^{+0.07}_{-0.11}$ GeV
s	-1/3	1/2	104^{+26}_{-34} MeV
t	2/3	1/2	172.2 ± 2.1 GeV
b	-1/3	1/2	$4.2^{+0.17}_{-0.07}$ GeV
e	-1	1/2	$0.511 \pm 1.3 \times 10^{-8}$ MeV
μ	-1	1/2	$105.658 \pm 4.0 \times 10^{-6}$ MeV
τ	-1	1/2	1776.84 ± 0.17 MeV
ν_e	0	0	< 2 eV
ν_ν	0	0	< 0.19 MeV
ν_τ	0	0	< 18.2 MeV
γ	$< 5 \times 10^{-30}$	1	1×10^{-18} eV
W^-	-1	1	80.40 ± 0.025 GeV
W^+	1	1	80.40 ± 0.025 GeV
Z^0	0	1	91.19 ± 0.0021 GeV
$g^i, i = 1..8$	0	1	0 (Assumed)

the group $SU(2)$ (Special Unitary group of order 2). In general the $SU(2)$ group is represented by 2×2 matrices which describe rotations of the fields in a two dimensional complex space (*i.e* spin rotations). Finally, a third internal symmetry should be considered for all particles: $SU(3)$. The SU(3) group is represented by 3×3 matrices which describe rotations of the fields in a three dimensional complex space.

In the SM, the electromagnetic and weak interaction are unified in a single interaction known as the electroweak interaction. This unification is based on postulating that the electroweak interaction must satisfy a local gauge symmetry, i.e. a situation where the physics must be invariant under a transformation that adds a phase to the fields, this phase being different in every point in space time. According to the theory, the four carriers of the interaction must be massless which is in disagreement with the experimental evidence. This fact indicates that the symmetry is broken in nature in the group $SU(2) \times U(1)$. However, the SM postulates that the symmetry

is not broken in the interaction itself but rather in the ground state, or vacuum. This is known as spontaneous symmetry breaking [7]. The spontaneous symmetry breaking is a fundamental concept on which the Higgs Mechanism is based.

In the next section, we will introduce the initial attempts to describe the kinematics of particles, which sets the basis for the Higgs Mechanism.

II.2.1 The Klein-Gordon and the Dirac Equations

Oskar Klein and Walter Gordon proposed an equation to describe the motion of free particles relativistically:

$$H = p^2 + m^2 \Rightarrow (\partial_\mu \partial^\mu - m^2)\phi(x) = 0 \tag{2}$$

In Equation 2, H represents the Hamiltonian (energy), p the momentum and m the mass of the particles. Nevertheless, this equation predicted the possibility of solutions with negative energies and probabilities for free particles. In order solve this "problem", the English physicist Paul Adrien Maurice Dirac, proposed a solution linearizing the momentum and mass of the particles in the Klein-Gordon equation by introducing some extra factors:

$$H = \gamma^0(\gamma \cdot p + m) \Rightarrow (i\gamma^\mu \partial_\mu - m)\psi(x) = 0 \tag{3}$$

where ψ represents the wave function that describes a particular fermion. The γ^μ factor in the Dirac equation, represents a set of 4×4 matrices, that are related to spin rotations in the $SU(2)$ space [8].

Despite Dirac's efforts, his results showed solutions where particles could have negative energies, as in the Klein-Gordon equation, but probabilities were always positive. Later it was understood that the negative solutions represented the possibility of anti-particles, which indeed were experimentally found.

Now we understand that the Klein-Gordon equation describes the relativistic motion of free bosons and the Dirac formalism does it for fermions.

We can summarize the kinematics of all fermions in the Dirac formulism in terms of a lagrangian function [1].

$$\mathcal{L}_{ferm} = \sum_{\psi} \overline{\psi} i\gamma^{\mu} D_{\mu}\psi + m\overline{\psi}\psi \tag{4}$$

Where we have used the covariant derivative D_{μ} defined in Equation 5, instead of the partial derivative ∂_{μ} initially introduced in the Dirac equation.

$$D_{\mu} = \partial_{\mu} - ig_1\frac{Y}{2}B_{\mu} - ig_2\frac{\tau^i}{2}W^i_{\mu} - ig_3\frac{\lambda^a}{2}G^a_{\mu} \tag{5}$$

In Equation 5, B_{μ}, W^i_{μ} and G^a_{μ} are spin-one fields needed to keep the gauge invariance under $U(1)$, $SU(2)$ and $SU(3)$ transformations respectively. The terms Y, τ^i and λ^a are the generators of transformations for each individual group. The strength of the interactions is carried by the coupling constants g_1, g_2 and g_3.

Even though the Dirac Lagrangian is an elegant and simple way to describe the kinematics of fermions, the mass term is not invariant under $SU(2)$ transformations. In particle physics, the projection of the spin direction onto the momentum of the particle is known as helicity. Helicity for neutrinos is not conserved under a parity exchange. Experimentally, we have not observed neutrinos with their momentum and spin projections pointing in the same direction, which is referred to as right handed helicity. This experimental fact is introduced into the theory by requiring that left-handed fermions must be put into $SU(2)$ doublets and right handed fermions into $SU(2)$ singlets (no right handed neutrinos!):

$$\begin{pmatrix}\nu_e \\ e\end{pmatrix}_L \ , \ \begin{pmatrix}\nu_\mu \\ \mu\end{pmatrix}_L \ , \ \begin{pmatrix}\nu_\tau \\ \tau\end{pmatrix}_L \ , \ e_R \ , \ \mu_R \ , \ \tau_R \tag{6}$$

For this reason in Equation 4, the mass term can be expressed as:

$$m\overline{\psi}\psi = m(\overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R) \tag{7}$$

The product of $\overline{\psi}_R\psi_L$ and $\overline{\psi}_L\psi_R$ are not $SU(2)$ singlets and would not give an invariant Lagrangian under $SU(2)$ transformations. The only way to preserve the gauge invariance of the Lagrangian is to set the mass to zero for all particles, which is clearly against the experimental evidence. Peter Higgs, an English theoretical physicist, proposed a solution for this problem.

II.2.2 Higgs Mechanism

The standard model includes an additional interaction between all particles due to a still unobserved field known as the Higgs field. When this additional interaction is added to the theory, it is possible to accommodate intermediaries with mass which solves the inconsistency of the massless carriers in the initial approach of the electroweak unification. In addition, the matter of all the particles in the SM springs essentially from the Higgs field interaction.

In the SM, the Lagrangian density for the Higgs field is described as follows:

$$\mathcal{L}_{higgs} = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{8}$$

where the first term in Equation 8 corresponds to the dynamical interactions of the Higgs field and the last two terms represent the Higgs potential ($V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$), with μ^2 and λ initially introduced as free parameters.

This potential is symmetric under phase changes that transform the field from $\phi \Rightarrow -\phi$. Detailed studies have shown that terms higher than fourth order for the field lead to infinities. In order to keep the potential bounded from below when $\phi \rightarrow \infty$ we must require that λ must be greater than zero.

The standard model Higgs field is expressed as a doublet in terms of two complex fields:

$$\phi = \begin{pmatrix} \phi^\dagger \\ \phi \end{pmatrix} \quad \text{where} \quad \phi^\dagger = \frac{\phi_1 + i\phi_2}{\sqrt{2}} \quad \text{and} \quad \phi = \frac{\phi_3 + i\phi_4}{\sqrt{2}} \tag{9}$$

In order to find the ground state of the field, we should calculate the minimum of the Lagrangian. We can minimize the kinetic energy by considering the field as constant. So, minimizing the potential we get:

$$\phi^\dagger\phi = -\frac{\mu^2}{2\lambda} = \frac{\nu^2}{2} \quad \text{where we defined} \quad \nu^2 = -\frac{\mu^2}{\lambda} \tag{10}$$

Using the definitions of the Higgs doublets from Equation 9, we find that $\phi^\dagger\phi$ is defined as:

$$\phi^\dagger\phi = \frac{\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2}{2} \tag{11}$$

As we have seen, the SM Higgs field has four components which don't have a graphical representation. In order to get an idea of the shape of the ground state, we can use a representation of the Higgs field in terms of just two components (one scalar and and the other imaginary), this is known as the Abelian Higgs field [1]. In Figures 1 and 2 we show the graphical representation of the potential for the Abelian Higgs field.

For the case where $\mu^2 < 0$ the potential happens to have two stable minima in the ground state around $\pm\nu^2$. If we choose $+\nu$ or $-\nu$ as our ground state, we should choose a specific direction of the field. Using Equation 11

$$\phi_3 = \nu, \quad \phi_1 = \phi_2 = \phi_4 = 0 \quad \rightarrow \quad \phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{12}$$

By expanding around this minimum, we can find the excited states of the field which according to quantum field theory are the particles. If the perturbation is $H(x)$, we get:

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix} \tag{13}$$

Because a specific direction has been chosen, we have broken the original symmetry of the system. By using this principle and introducing a covariant derivative in

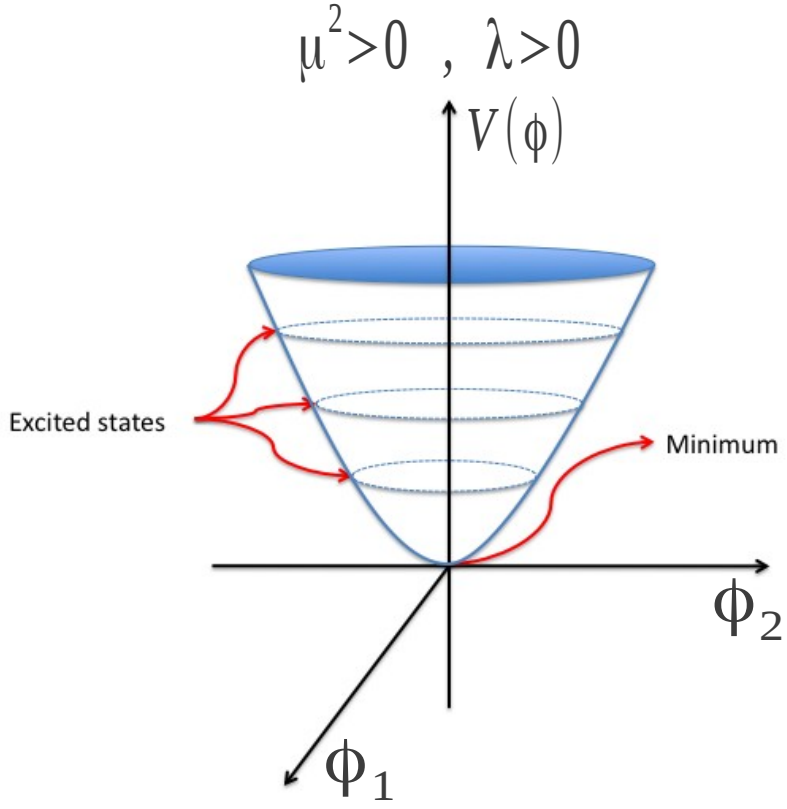


Fig. 1. Representation of the Abelian Higgs potential in the standard model. This potential corresponds to the case where the parameters for the Higgs Lagrangian introduced in Equation 8 have the values of $\mu^2 > 0$ and $\lambda > 0$.

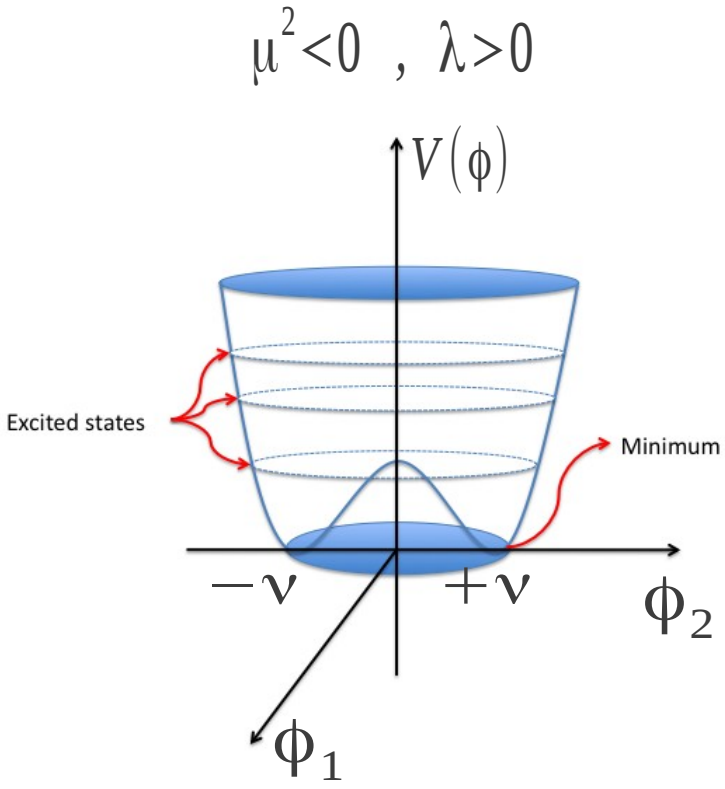


Fig. 2. Representation of the Abelian Higgs potential in the standard model. This potential corresponds to the case where the parameters for the Higgs Lagrangian introduced in Equation 8 have the values of $\mu^2 < 0$ and $\lambda > 0$.

Equation 8 to preserve the invariance under $U(1) \times SU(2)$ transformations (Equation 14), it is possible to accommodate in the electroweak unification three massive gauge bosons and a massless one (Goldstone boson), and agree with the experimental evidence.

$$D_\mu = \partial_\mu - ig_1 \frac{Y}{2} B_\mu - ig_2 \frac{\tau^i}{2} W_\mu^i \tag{14}$$

Including the extra terms in the expression for the covariant derivative in Equation 14, Equation 10 becomes:

$$\begin{gathered} \phi^\dagger \left(ig_1 \tfrac{Y}{2} B_\mu + ig_2 \tfrac{\tau^i}{2} W_\mu^i \right)^\dagger \left(ig_1 \tfrac{Y}{2} B^\mu + ig_2 \tfrac{\tau^i}{2} W^{i\mu} \right) \phi i = \\ \tfrac{1}{8}\mu^2 g_2^2 \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) + \tfrac{1}{8}\mu^2 \left(g_1 B_\mu - g_2 W_\mu^3 \right) \end{gathered} \tag{15}$$

By defining:

$$W^+ = \frac{-W^1 + iW^2}{\sqrt{2}} \quad , \quad W^- = \frac{-W^1 - iW^2}{\sqrt{2}} \qquad W^0 = W^3 \tag{16}$$

we can express the first term after the equal sign in Equation 15 as:

$$\frac{1}{8}\mu^2 g_2^2 \left((W_\mu^1)^2 + (W_\mu^2)^2 \right) = \left(\frac{1}{2} g_2 \right)^2 W_\mu^+ W^{-\mu} \tag{17}$$

where the term $\frac{1}{2}g_2$ represents the mass of the electroweak W boson ($M_W = \frac{1}{2}g_2$). Similarly, for the second term in Equation 15 we obtain:

$$\frac{1}{8}\mu^2 \left(g_1 B_\mu - g_2 W_\mu^3 \right) = \frac{1}{8}\nu^2 \left(g_1^2 + g_2^2 \right) Z_\mu Z^\mu \tag{18}$$

which gives the mass of the neutral electroweak boson (Z), defined as $M_Z = \frac{1}{2}\sqrt{g_1^2 + g_2^2}$. In Equation 18, the Z_μ field is defined as:

$$Z_\mu = \frac{g_1 Y_L B_\mu + g_2 W_\mu^0}{\sqrt{g_2^2 + g_1^2 Y_L^2}} \tag{19}$$

A great success of this theory was the accurate prediction of the mass of W and the Z bosons before they were experimentally found, which guides us to think that the Higgs mechanism is possibly a correct approach to explain how particles acquire mass.

II.2.3 Feynman Diagrams

Feynman diagrams are visual representations of radiative and/or interaction processes of particles. Each component of the diagram is related with a mathematical term that describes the nature of the interaction. Figure 3 shows a Feynman diagram for a tau that decays leptonically.

The solid arrows in Figure 3, represent particles in the initial and final state of the specific physics process. These arrows, directed to or from vertices, are related to the direction of time and have a different meaning for particles and antiparticles. For particles, incoming lines to a vertex represent the initial state and ongoing lines from the vertex represent the final states. In this thesis antiparticles in final states are represented with lines pointing towards the vertex and the propagator of the interaction is represented by wavy lines connecting vertices.

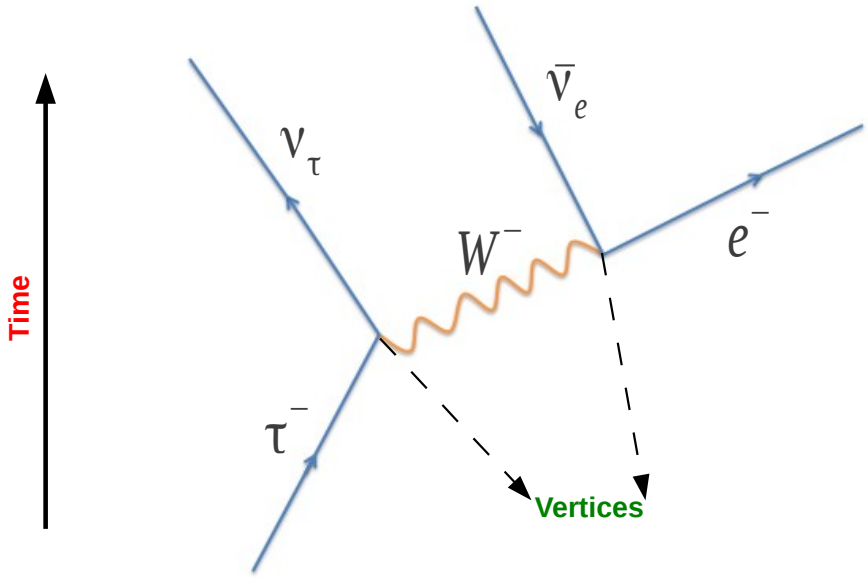


Fig. 3. Pictorical representation of the leptonic decay of a tau, known as Feynman diagram. Each part of the diagram is related with a mathematical term that describes the physics process.

CHAPTER III

Z' MODELS

The Z' gauge boson appears in several extensions of the standard model. After the introduction of the electroweak unification described by the symmetry groups $SU(2) \times U(1)$, theories containing the addition of new symmetry groups, motivated by the called Grand Unified Theories (GUT) were proposed. These theories postulate the unification for all the known interactions, at a large energy scale, into a single interaction responsible of all the processes among particles in the early Universe. Models based on E_6 symmetry groups [9], little Higgs, extra dimensions etc. have a Z' boson at different energy scales. The Z' naturally arises in supersymmetry and string theories. The most simple extension of the known symmetry groups involves the addition of $U(1)'$ factors in alternative electroweak gauge theories [3].

If the Z' is experimentally found it will have a dramatic impact in modern particle physics. For instance, the existence of the Z' is directly attached with new symmetry groups, for example $U(1)'$, which leads to new particles. Also, if the Higgs field is responsible for the mass generation of particles, the Z' will acquire mass through the symmetry breaking of the new symmetry groups (*i.e.* $U(1)'$) with the $SU(2)$ SM groups, which requires an extended Higgs sector [3].

III.1 Z' Mass in the Extended $SU(2) \times U(1) \times U(1)'$ Group

The Z' is defined as a gauge boson with neutral electric charge. In the standard model, the photon and the Z boson are the mediators of the neutral interactions among fermions. The physics processes that govern these neutral interactions are described with the following Lagrangian:

$$-\mathcal{L}_{nc}^{SM} = eJ_{em}^{\mu}A_{\mu} + g_1 J_1^{\mu} Z_{1\mu}^0 \tag{20}$$

In Equation 20 g and g' are the couplings for the weak and electromagnetic interactions, J^{μ}_{em} and J^{μ}_{1} are the electromagnetic and weak currents among fermions and, A_{μ} and $Z^{0}_{1\mu}$ are the boson fields needed to keep the gauge invariance as explained in Section II.2. The fields are defined as:

$$\begin{aligned} A_{\mu} &= sin\theta_{W} W_{3\mu} + cos\theta_{W} B_{\mu} \\ Z^{0}_{1\mu} &= cos\theta_{W} W_{3\mu} - sin\theta_{W} B_{\mu} \end{aligned} \tag{21}$$

where $\theta_{W} = tan^{-1}(g^{'}/g)$ is known as the weak angle [1].

When the $U(1)'$ extra symmetry is introduced, a similar Lagrangian describes the neutral currents of fermions:

$$-\mathcal{L}^{ext}_{nc} = eJ^{\mu}_{em} A_{\mu} + \sum_{\alpha=1}^{n+1} g_{\alpha} J^{\alpha}_{\mu} Z^{0}_{\alpha\mu} \tag{22}$$

This Lagrangian includes extra terms for the couplings g_{α}, the neutral weak current J^{α}_{μ} and the boson field $Z^{0}_{\alpha\mu}$, in order to account for the additional $U(1)'$ symmetry groups. When α is equal to one, we obtain the SM Lagrangian for neutral interactions among fermions shown in Equation 20.

The mass generation of the Z' for this model is based on the Higgs mechanism. In an analogous way with the SM Higgs, the Z' acquires mass through the symmetry breaking of the $SU(2) \times U(1) \times U(1)'$ group. As explained in Section II.2.2 the kinematic interactions of a scalar field are described by the Lagrangian:

$$\mathcal{L}^{kinetic} = (D^{\mu}\phi_{i})^{\dagger} D_{\mu}\phi_{i} \tag{23}$$

where the covariant derivative is defined as:

$$D_{\mu}\phi_{i} = \left(\partial_{\mu} + ieq_{i}A_{\mu} + i\sum_{\alpha=1}^{n+1} g_{\alpha} Q_{\alpha i} Z^{0}_{\alpha\mu} \right) \phi_{i} \tag{24}$$

In Equation 24 q_{i} is the electric charge of fermions and $Q_{\alpha i}$ is the charge associated with the $U(1)_{\alpha}$ symmetry group. The $Q_{\alpha i}$ term represents the couplings for fermions that account for the left handed and right handed symmetries (see Section II.2).

If we choose a particular vacuum expectation value, we obtain a mass term for the $Z^0_{\alpha\mu}$ field while A_μ remains massless (photon field):

$$\begin{aligned} \mathcal{L}_{Z_{mass}} &= \tfrac{1}{2} M^2_{\alpha\beta} Z^0_{\alpha\mu} Z^{0\mu}_{\beta} \\ M^2_{\alpha\beta} &= 2 g_\alpha g_\beta \textstyle\sum_i Q_{\alpha i} Q_{\beta i} |\langle \phi_i \rangle|^2 \end{aligned} \tag{25}$$

The mass expression in Equation 25 can be represented as a matrix with multiple terms depending on the different possible combinations of α and β. By diagonalizing the elements of the mass matrix, we can obtain multiple eigenstates for the vector field $Z_{\mu\alpha}$ [3]:

$$Z_{\alpha\mu} = \sum_{\beta=1}^{n+1} U_{\alpha\beta} Z^0_{\beta\mu} \tag{26}$$

where $U_{\beta\alpha}$ is an orthogonal matrix that accounts for the possible mixed states depending on the values of α and β. The simplest case is when n is equal to one, which yields the square mass mixed matrix, as defined in Equation 25, for the Z and Z' bosons:

$$M^2_{Z-Z'} = \begin{pmatrix} 2 g_1^2 \sum_i t^2_{3i} \langle \phi_i \rangle^2 & 2 g_1 g_2 \sum_i t^2_{3i} Q_i \langle \phi_i \rangle^2 \\ 2 g_1 g_2 \sum_i t^2_{3i} Q_i \langle \phi_i \rangle^2 & 2 g_2^2 \sum_i Q_i^2 \langle \phi_i \rangle^2 \end{pmatrix} \tag{27}$$

where t_{3i} is the third component of weak isospin that, combined with the weak hypercharge, gives rise to the definition of electric charge in the SM [1]:

$$t_{3i_L} + y_{i_L} = t_{3i_R} + y_{i_R} = q_i \tag{28}$$

The square mass matrix defined in Equation 27 can be expressed as:

$$M^2_{Z-Z'} = \begin{pmatrix} M^2_{Z^0} & \Delta^2 \\ \Delta^2 & M^2_{Z'} \end{pmatrix} \tag{29}$$

where Δ represents mixed mass states of the SM Z boson and the Z'.

CHAPTER IV

THE LHC

The Large Hadron Collider is a proton-proton particle accelerator built at the European Center for Nuclear Research (CERN), located at the border between France and Switzerland (near Geneva Switzerland). The LHC also collides heavy ions to study fundamental physics, especially high energy states where quarks and gluons behave almost as free particles forming a quark-gluon plasma. Nevertheless, the topics of this thesis will focus on the physics of proton-proton collisions and not heavy ions.

Built inside an underground tunnel (from 50 to 150 m deep), as shown in Figure 4, with a perimeter of 27 km, the LHC has been designed to achieve the highest ever man-made collision energy. The LHC has a total of 1232 superconducting dipole magnets to guide the beam through the ring. Using liquid helium, the magnets are cooled down to a temperature of 2.0 K to keep their superconducting state. This extreme low temperature helps as well to optimize the vacuum inside the beam pipe.

The protons are conducted through a set of linear and circular accelerators (synchrotrons), where their speed increases and are clustered in bunches that have an average separation of around 50 ns (50 ns at the moment, 25 ns at designed parameters) between them. Initially, the protons are produced and conducted to a linear accelerator called LINAC2, which increases their energy to 50 MeV. Then, the acceleration of the protons is sequentially increased with a set of synchrotrons before they are transmitted to the LHC. The synchrotrons use radio frequency cavities (r.f) to boost the protons. After the LINAC2, the protons are injected to the Proton Synchrotron Booster (PSB) where they reach an energy of 1.4 GeV. Once the protons leave the PSB they are transferred to a larger proton synchrotron (PS) where their energy is increased up to 25 GeV. After the PS, the bunches are accelerated to 450 GeV in a Super Proton Synchrotron (SPS) and then are transferred to the LHC where they finally reach their maximum speed (close to the speed of light

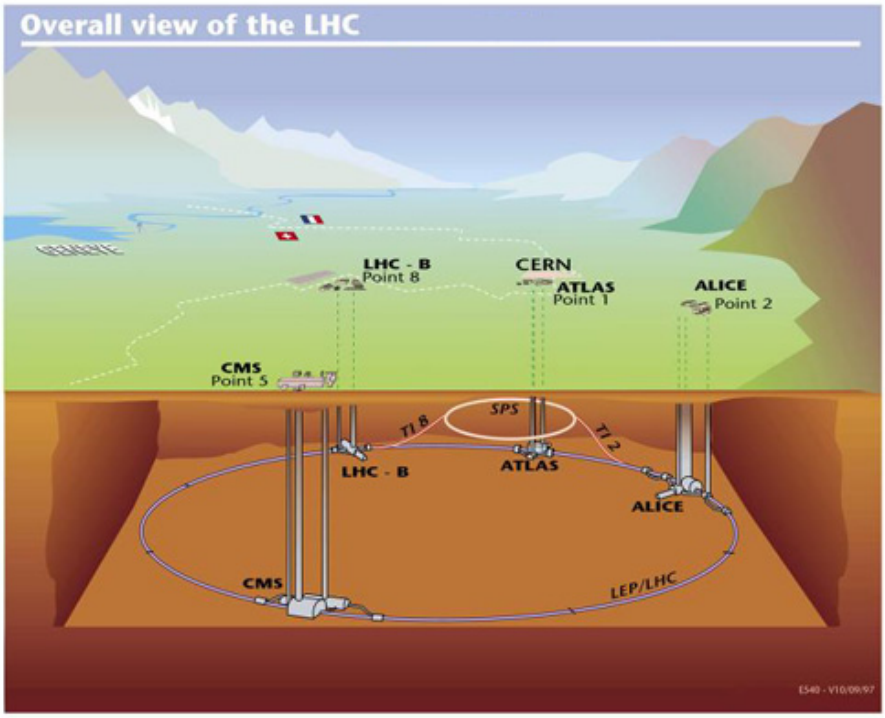


Fig. 4. Illustration of the LHC, built inside an underground tunnel located at the border between France and Switzerland.

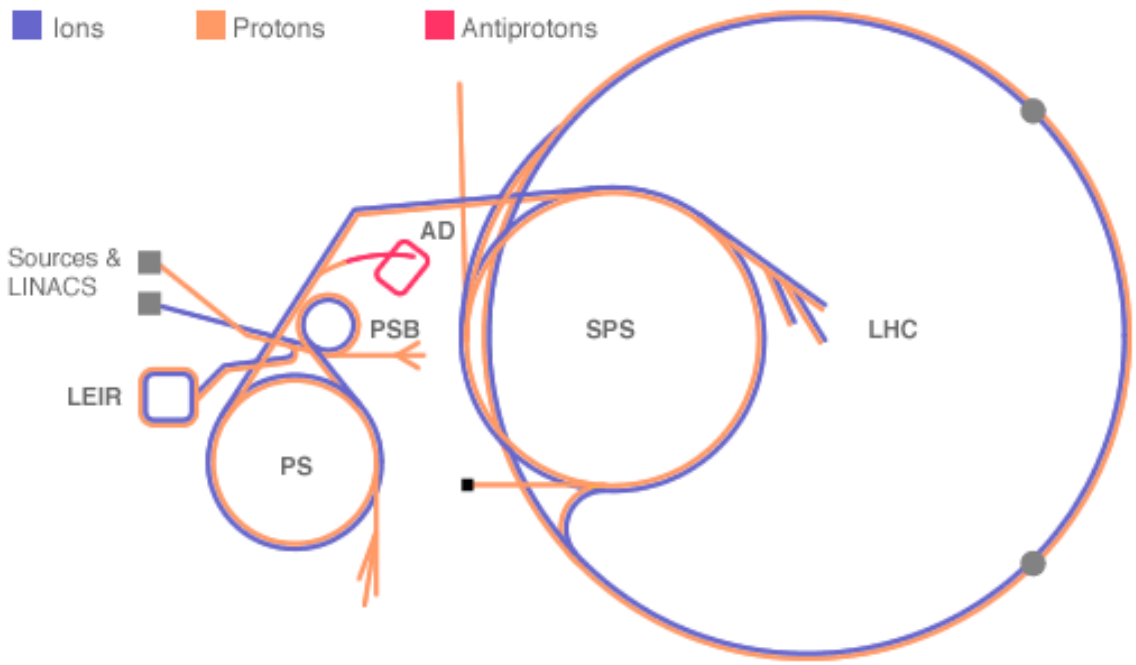


Fig. 5. Illustration of the accelerator complex at CERN where the LHC is located. The PS and SPS synchrotrons pre-accelerate the protons before they enter to the LHC.

0.999999c) and collide. Nominally, protons will reach a collision energy of 7 TeV (14 TeV at the center of mass). In 2010 the LHC achieved a collision energy of 3.5 TeV per beam. Figure 5 illustrates the accelerator complex just described.

Luminosity is an important quantity for particle accelerators and experiments; it allows us to estimate the rate of collisions per bunch crossing and thus determine the accelerator performance. In order to have a higher luminosity we can increase the number of particles per bunch and focus the beam to have a smaller cross section which improves the probability of having a hard collision. A higher rate of bunches that can collide is another way to increase the luminosity which can be complementary to the number of particles per bunch. Equation 30 shows the expression to estimate the luminosity in a particle collider.

$$L = \frac{N_b^2 n_b f \gamma_r}{4\pi\varepsilon_n \beta^*} F \tag{30}$$

In Equation 30, N_b represents the number of protons per bunch, n_b is the number of colliding bunches per beam, f is the frequency of revolution, γ is the relativistic Lorentz factor, ε_n is the transverse beam emittance, β^* is the beta function at the center of collision and F is a geometrical factor related to the crossing angle of the beams [10].

Luminosity is measured using the inverse of a unit known as a barn (b). One barn is equivalent to 10^{-24} cm^2. The instantaneous luminosity is measured in terms of $b^{-1}s^{-1}$. The total luminosity delivered by the accelerator corresponds to the integral of the instantaneous luminosity in a time range. A measure of the likelihood of a physics process to occur is referred to as a cross-section. The total standard model cross section at nominal LHC energies is approximately 100 mb. Therefore, under optimal operation the LHC is expected to reach a nominal luminosity of $L = 10^{34} cm^{-2} s^{-1}$ which leads to 1 billion interactions per second.

IV.1 LHC Physics Goals

The LHC has a total of six experiments that will study diverse topics in high energy physics. Four experiments are located at the LHC collision points: CMS, ATLAS, LHCb (Large Hadron Collider Beauty) and ALICE (A Large (Heavy) Ion Collider Experiment) and two experiments are located near the collision points: TOTEM (TOTal Elastic and diffractive cross section Measurement) and LHCf (Large Hadron Collider forward), close to CMS and ATLAS respectively. CMS and ATLAS are the largest experiments at the LHC and both pursue the same physics research. One of the main goals of these two experiments is to elucidate the spontaneous symmetry breaking in the electroweak interaction and determine if the Higgs field is responsible for it. There are other important questions to be answered too such as: what is the nature of dark matter? Does the universe have extra-dimensions? Is Supersymmetry true? etc.

LHCb is an experiment focused on the study of bottom quark physics (b physics). The study of CP-violation using b-hadrons (particles that contain a b quark) is one of the main objectives of this experiment. ALICE is a heavy ion experiment looking to study heavy nuclei collisions. Their research is mainly focused on the study of strong interactions in matter at high densities, specifically states where quark-gluon plasmas are expected to form.

LHCf will use the particles coming out the ATLAS interaction point in the forward direction. The objective of this experiment is to have a better understanding of the cascades that cosmic rays generate when they collide with the atmosphere of the earth. This experiment will simulate these cascades, using particles that come in the forward direction from the collision point. It is expected that the better understanding of these cascades will allow scientists working in large cosmic ray experiments to improve their analyses and contribute to the design of future detectors.

TOTEM will use, as a source, particles coming from the forward direction of the interaction point where CMS is located. The aim of this experiment is to measure the total cross section of proton-proton collisions at the LHC.

CHAPTER V

THE CMS DETECTOR

With a total weight of 12,500 tons, a length of 21.6 m and a diameter of 14.6 m the CMS detector is one of the two largest detectors at the LHC (Figure 6). The detector consists of several sub-detectors enclosed in different layers: Silicon Pixels and Silicon Strips for tracking reconstruction, Electromagnetic and Hadronic Calorimeters for energy measurements and three different kinds of muon detectors. The tracker system, the electromagnetic calorimeter and a section of the hadronic calorimeter are enclosed in a superconducting solenoid magnet.

Figure 7 shows a slice of the CMS detector cross section. The figure illustrates signatures left by different kinds of particles in the sub-detectors. Particles that have an electric charge, such as electrons, muons and charged hadrons, leave a signature in the tracker system and bend in the presence of the magnetic field. By measuring the radius of curvature and the direction of the velocity of the particle with respect to the magnetic field, we can find the momentum of the particle using Lorentz force equation.

$$mv = qRBsin\theta \tag{31}$$

where q is the electric charge, R is the radius of curvature, B is the magnetic field and θ is the angle between the velocity of the particle and the magnetic field.

Electrons and photons deposit most of their energy in the electromagnetic calorimeter through the generation of cascades of particles also known as showers. Charged pions have dominantly strong interactions and deposit most of their energy in the hadronic calorimeter.

Muons have a mass around two hundred times bigger than the electron. For this reason the momentum of a muon is two hundred times bigger than the momentum of an electron that moves with the same velocity. Because of this muons do not bend as much for the same velocity as electrons in the magnetic field. As consequence the

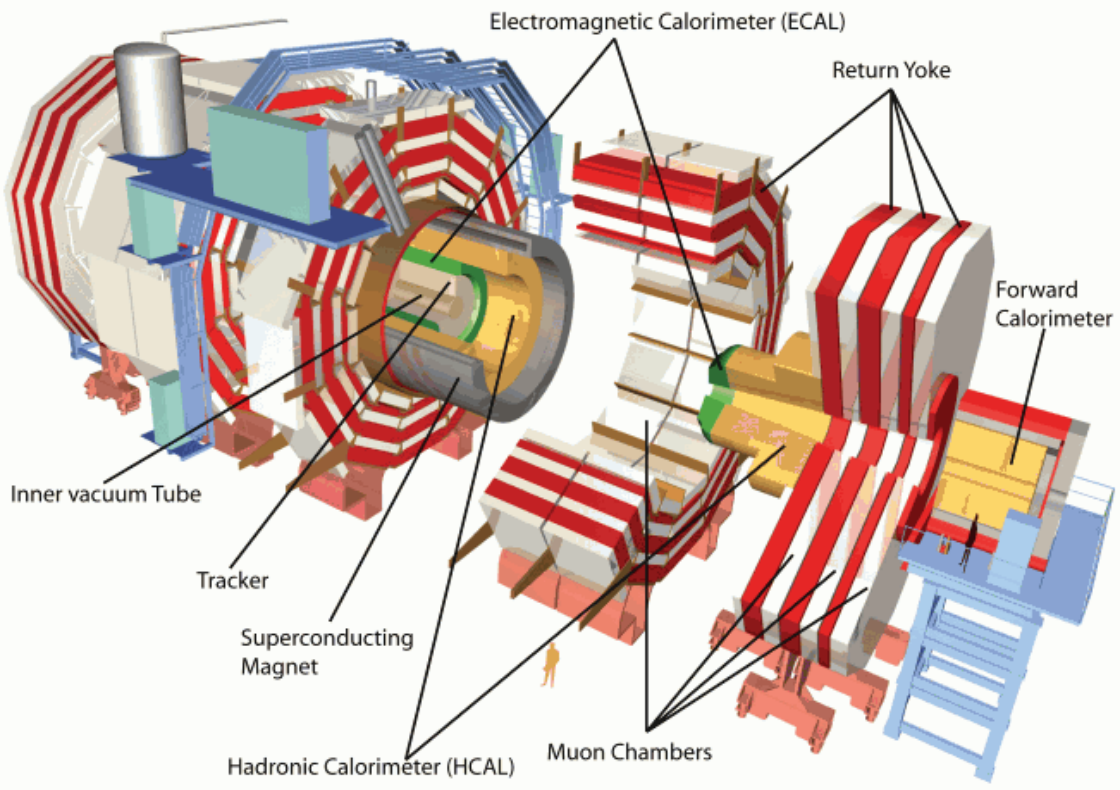


Fig. 6. CMS detector diagram. The detector is composed of several sub-detectors used to measure the position, momentum and energy of the particles resulting from high energy collisions. The detectors use different technologies in order to identify different signatures left by the particles.

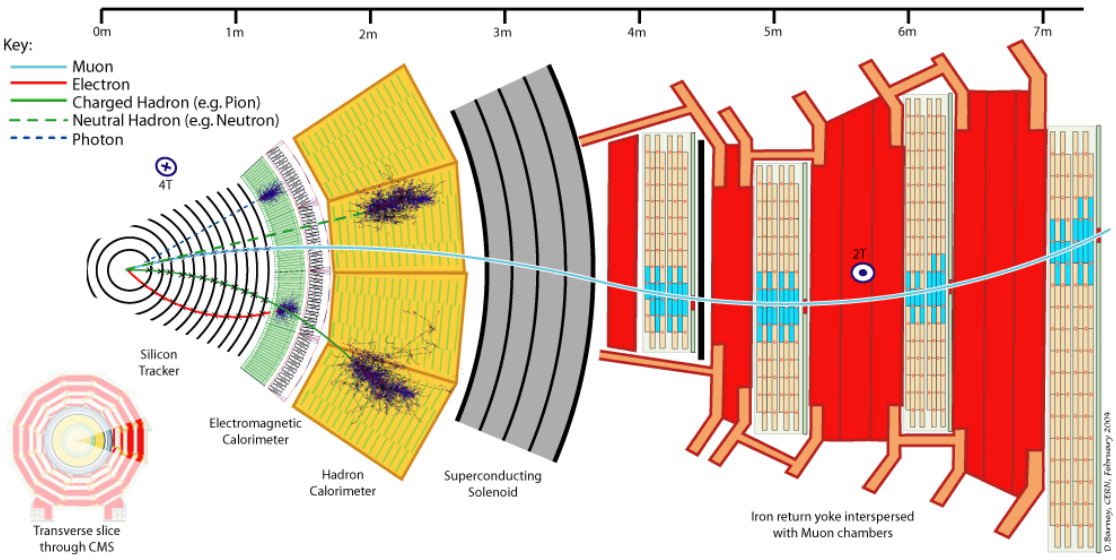


Fig. 7. CMS detector slice. The figure shows the signatures of different particles in the CMS sub-detectors. The tracker is used to reconstruct the trajectory of charged particles. Charged particles bend in the presence of a strong magnetic field produced by a superconducting solenoid. The curvature of the charged particles measured in the tracker material is used to estimate momentum. The energy of electrons and photons is measured in the electromagnetic calorimeter. The hadronic calorimeter is used to measure the energy of particles that have strong interactions. The outer layers of the CMS detector have a set of muon detectors in order to measure the momentum and position of these particles.

momentum of muons with high velocity can not be accurately measured in the tracker system. Furthermore, the small change in the direction of the muon acceleration results in low bremsstrahlung radiation. Because of this, the energy deposition of muons in the electromagnetic calorimeter is very small. The different muon detectors in the outer layers of the the CMS detector have been designed to measure more accurately the position and momentum of muons.

V.1 The CMS Coordinate System

CMS uses a right handed coordinate system. The positive Z axis points in the anticlockwise-beam direction. The x axis points to the center of the LHC ring, and the Y axis points up, perpendicular to the LHC plane. Spherical coordinates are used to describe the position of the detector components and reconstructed particles. The polar angle, θ, is measured with respect to the Z axis and the azimuthal angle, ϕ, is measured with respect the $X - Y$ plane as shown in Figure 8.

A coordinate called pseudorapidity (η) is commonly used in particle physics instead of the polar angle. This new coordinate is defined as:

$$\eta = -ln\left(tan\left(\frac{\theta}{2}\right)\right) \tag{32}$$

The cross section for a head on collision between two protons is very small. Therefore, the probability to have this hard scattering interaction with high energy exchange is low. A high energy exchange increases the probability for the production of heavy particles. At the other extreme soft scattering processes have a large cross section. The soft scattering implies a small change in the initial direction of the momentum of the colliding particles. By the law of conservation of momentum, the decay products resulting from the soft scattering, if any, should have a high component of momentum in the same direction as the initial particles. Because of this the number of particles in the forward region of the detector acceptance is higher than in the central part.

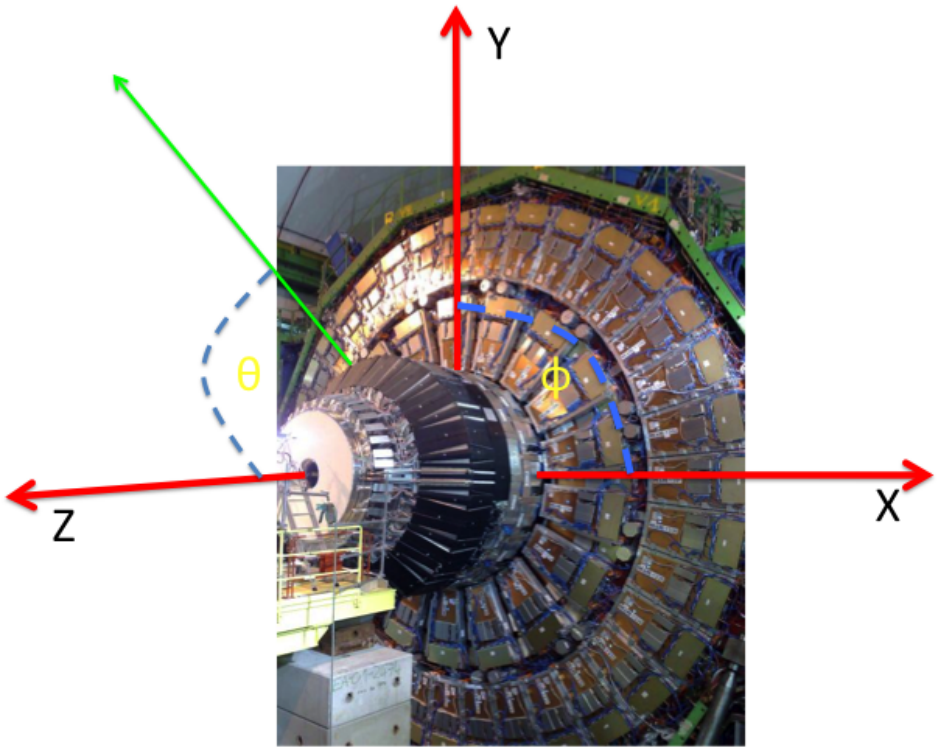


Fig. 8. Representation of the CMS coordinate system.

In order to obtain a more uniform distribution of particles in the detector acceptance across the polar coordinate, we use the η coordinate. To see the advantage of using η instead of θ, we can look at the rate of change of η with respect to θ:

$$d\eta = -\frac{d\theta}{sin\theta} \tag{33}$$

This gives a distribution in η with smaller regions in the forward direction and bigger regions in the central part of the detector, which results in an uniform distribution of particles across the polar coordinate.

V.2 Magnet Features

The magnetic field strength and its precise control are of fundamental importance for the CMS detector. With 2,168 turns, a length of 12.9 m and an inner bore of 5.9 m, the CMS magnet produces a magnetic field of 3.8 T which is required to perform a precise measurement of charged particle momenta. The most important features of the CMS magnet are listed in Table 2.

Table 2
Magnet Features.

Field	3.8 T
Inned bore	5.9 m
Length	12.9 m
Number of Turns	2168
Current	19.5 kA
Stored Energy	2.7 GJ

V.3 The Tracker System

The primary tracking system of CMS consists of two detectors: the Silicon Strips and the Pixel detector. The main objective of the tracking system is to provide an accurate measurement of the position and the electric charge of charged particles

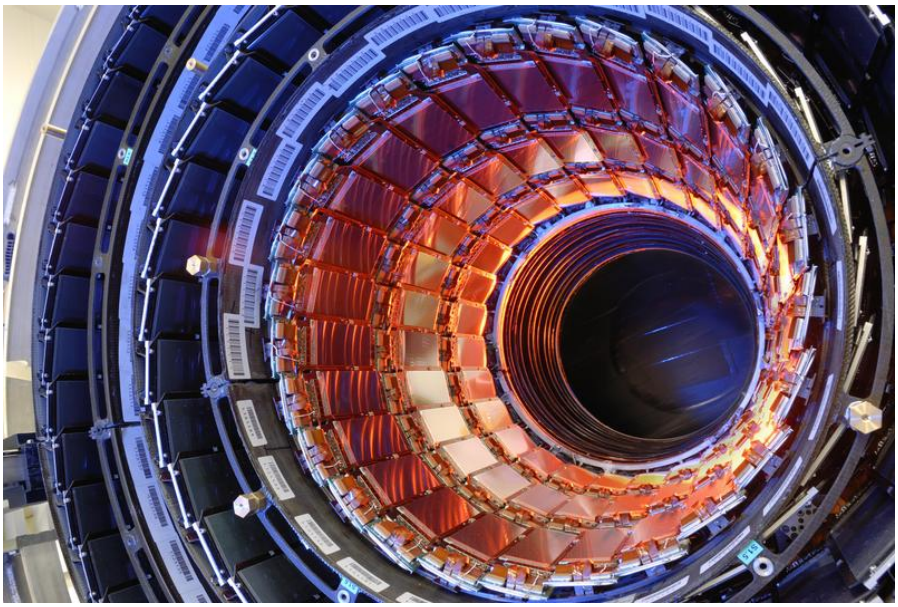


Fig. 9. Picture of the CMS tracker inner barrel (TIB) during the construction process. Picture taken from [11]

coming from a collision. At the LHC, one interesting topic of analysis will be reconstruction of $b - jets$ generated by heavy particle decays. To be able to identify this type of decay and other kinds of events, the tracking system should be close to the interaction point.

V.3.1 Silicon Strips

This system is composed of a total of 15,400 modules arranged in several barrel layers and endcaps. The barrel is divided into Tracker Inner Barrel (TIB) and Tracker Outer Barrel (TOB).

The TIB uses silicon sensors of 320 μm thickness and strip pitches of different width that vary from 80 to 120 μm. The TIB is organized in four cylindrical layers and is extended up to $|\eta| < 1.4$. Figure 9 shows a picture of the TIB during the construction process.

The TOB is composed of four identical disks, each with six cylindrical concentric layers. It covers a region up to $|\eta| < 1.8$. The TOB has silicon sensors of 500 μm thickness and strip pitches that vary from 80 to 120 μm.

A total of twelve disks on each side form the strip tracker endcaps. The disks are divided into nine disks called TEC (Tracker End Caps) and three smaller disks called TID (Tracker Inner Disks). The TID fill the gap between the TIB and the TEC. Silicon sensors with a thickness of 320 μm are used for the TID and the three innermost disks of the TEC, the remaining disks of the TEC have sensors of 500 μm. The Disks are concentrically arranged with respect to the beam pipe. The TID covers a pseudorapidity region of $1.2 < |\eta| < 2.1$ and the TEC covers a region of $1.0 < |\eta| < 2.5$.

Figure 10, shows a sketch of the TEC system on one side of the tracker barrel. A picture of one of the TEC systems is shown in Figure 11.

V.3.2 The Pixel Detector

The CMS Silicon Pixel Detector is the closest detector to the interaction point. It has about 66 million pixels, each one with a rectangular shape of 100×150 μm^2. The pixel detector provides two or more hits per track, allowing the detection of secondary vertices generated by relatively long-lived particles such as b or c quarks or tau leptons. Due to the high luminosity at the LHC, the pixel detector will be an important component of the CMS detector for reconstruction of events with many tracks.

The pixel detector consists of three barrel layers located at mean radii of 4.4, 7.3 and 10.2 cm, and two endcap disks (forward) assembled in a turbine-like geometry, extending from 6 to 15 cm in radius. In Figure 12 the CMS pixel detector geometry is shown. The three barrel layers have a length of 53 cm, and the endcap disks are placed at $|Z| = 34.5$ cm and 46.5 cm providing a coverage up to $\eta < 2.5$. The CMS pixel detector layers are composed of modular detector units of segmented

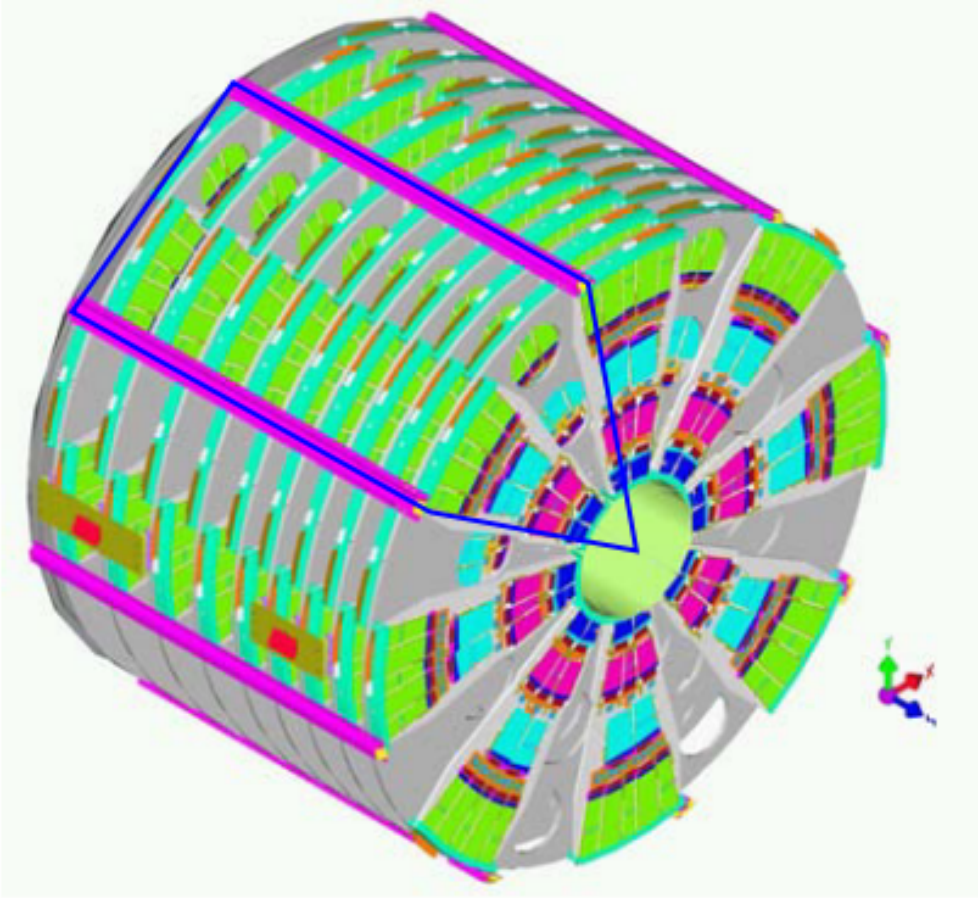


Fig. 10. Illustration of the TEC system on one side of tracker barrel. Picture taken from [11]

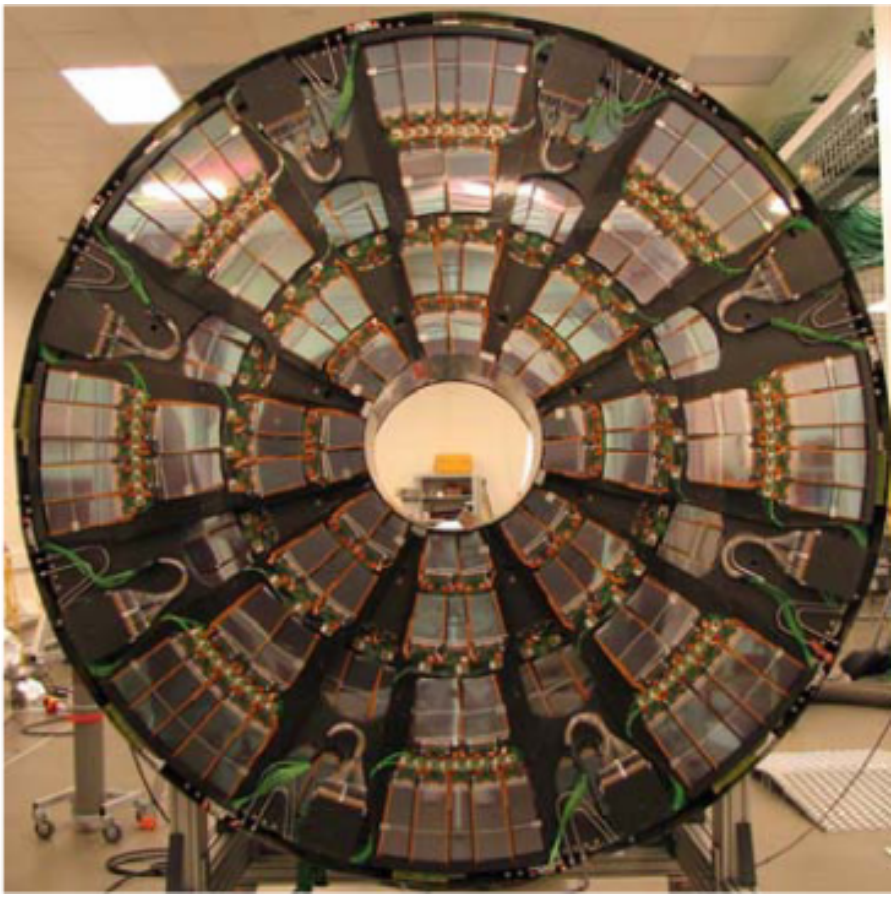


Fig. 11. Picture of one of the TEC systems. Picture taken from [11]

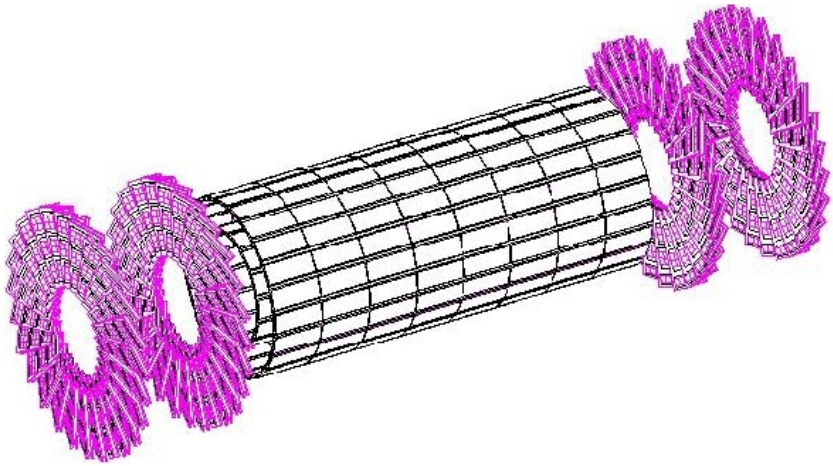


Fig. 12. Pixel Detector Geometry. The pixel detector is used for the reconstruction of primary and secondary vertices of charged particles.

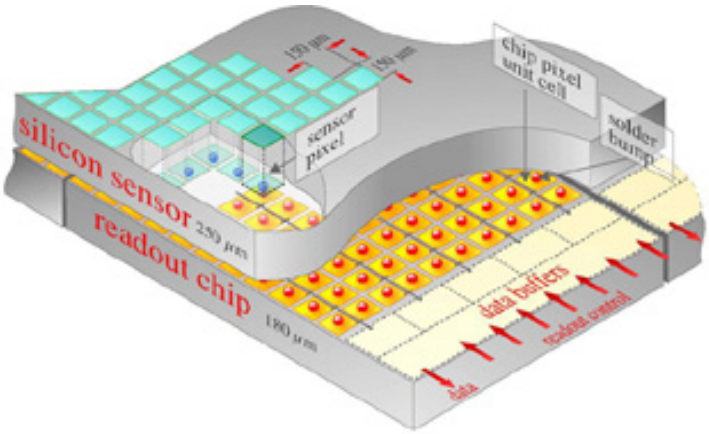


Fig. 13. Sketch of a pixel sensor. The sensor plates are connected to the readout chips using micro solder balls known as bump bonds.

sensor plates connected to a system of readout chips (ROC) using the bump bond technique (Figure 13).

V.4 The Electromagnetic Calorimeter

The CMS Electromagnetic Calorimeter (ECAL) is divided in two sections, Barrel ECAL (EB) and Endcaps ECAL (EE). This detector is used to measure the energy of electrons and photons. A total of 74,524 lead tungsten crystals ($PbWO_4$) are distributed between the EB and EE sections [11]. The crystals have a square cross section of $22 \times 22\ mm^2$ and a length of 230 mm each. When a charged particle passes through the ECAL crystals, it interacts with the material generating an electromagnetic shower. The photons then interact with the detector material producing electron positron pairs that radiate more photons developing the electromagnetic shower. The crystals are connected to a system of photo detectors that collect the scintillation light from the electromagnetic shower, amplify it and convert it to an electrical signal that is then sent to the readout system.

The EB has been structured in modules. A set of four modules, where the first one contains 500 crystals and the three remaining contain 400 each, compose a supermodule. Each supermodule covers half of the total barrel length. There are 36 supermodules distributed around the EB section which has a inner radius of 1.29 m and covers a pseudorapidity interval of $0 < \eta < 1.479$

Each EE is composed of two half disks, consisting of structural units of 5×5 crystals (supercrystals) of the same dimensions $24.7 \times 24.7 \times 220\ mm^3$. They are located at a distance of 314 cm from the interaction point and cover a pseudorapidity range of $1.479 < \eta < 3.0$. Figure 14 shows a sketch of the detector.

V.5 The Hadron Calorimeter

The HCAL has been designed to measure the energies and direction of motion of quarks, gluons and neutrinos, using the direction of jets (cones of hadrons and other particles) and the flow of missing transverse energy. The HCAL complements the electromagnetic calorimeter and the muon detector in the process of identification of photons, electrons and muons. The calorimeter is made of alternating layers of non-magnetic brass (alloy of copper and zinc) and stainless steel, to eliminate forces

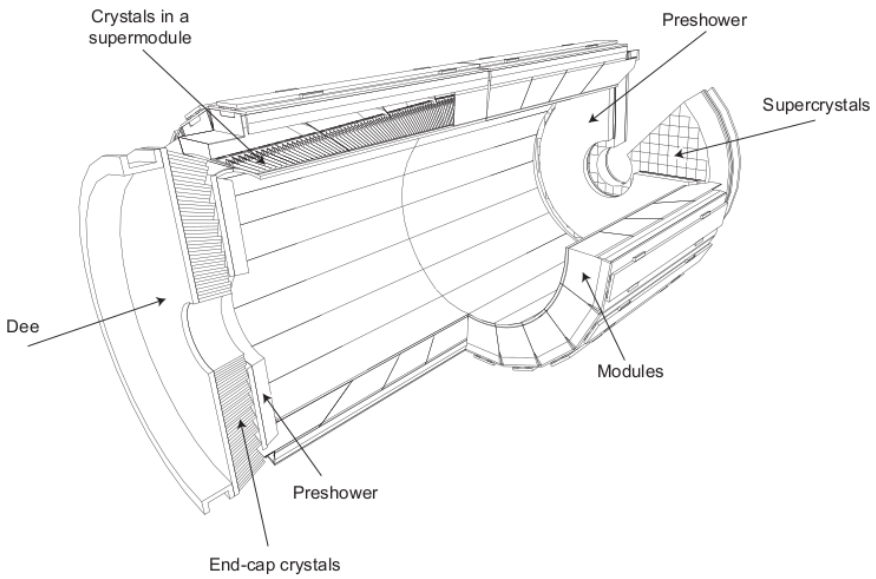


Fig. 14. Sketch of the electromagnetic calorimeter. This detector is used to collect the energy of particles that interact electromagnetically such as electrons and photons.

that can cause a distortion in the uniform internal magnetic field generated by the solenoid. Hadrons interacting with the detector material create a shower of particles. The shower is detected with plastic scintillators in between the metal layers and the light from the scintillators is sent to the readout system using optical fibers.

The HCAL is divided into three main segments: central, outer and forward. The central HCAL consists of the Hadron Barrel (HB) and the Hadron Endcap (HE). The HB is a cylindrical structure composed of two half barrels. Each half barrel contains a total of 16 towers, which cover a pseudorapidity region of $0 < \eta < 1.4$ giving a total of 1152 towers per half barrel with a ϕ segmentation of 5^o. The HE consists of sets of 14 towers covering a pseudorapidity region of $1.3 < \eta < 3.0$ with a segmentation of 5^o around ϕ for the 5 outermost towers (smaller η) and a segmentation of 10^o for the inner towers.

The Hadron Outer (HO) calorimeter is located outside the magnetic solenoid and physically inside the barrel muon system, covering a pseudorapidity region of $-1.26 < \eta < 1.26$. The HO is basically a layer of scintillators used to measure the energy of hadron showers that penetrate through the internal calorimeters, making it an energy tail-catcher [11].

The Hadron Forward (HF) calorimeter, covers a pseudorapidity region of $3 < |\eta| < 5$ and is located at 11.2 m from the interaction point. The HF contains steel/quartz fibers running parallel to the beam line. The signals captured in the fibers are channeled to photomultipliers. There are 13 towers in η with a segmentation of 10^o around ϕ except for the highest η one, where $\Delta\phi = 20^o$ which gives a total of 900 towers for the entire HF detector.

Figure 15 shows a sketch of a longitudinal view of the CMS detector, illustrating the main components of the HCAL detector.

V.6 The Muon System

Muons are one of the most interesting and important particles for CMS. It has been theoretically estimated that one of the golden modes for the Higgs decay is

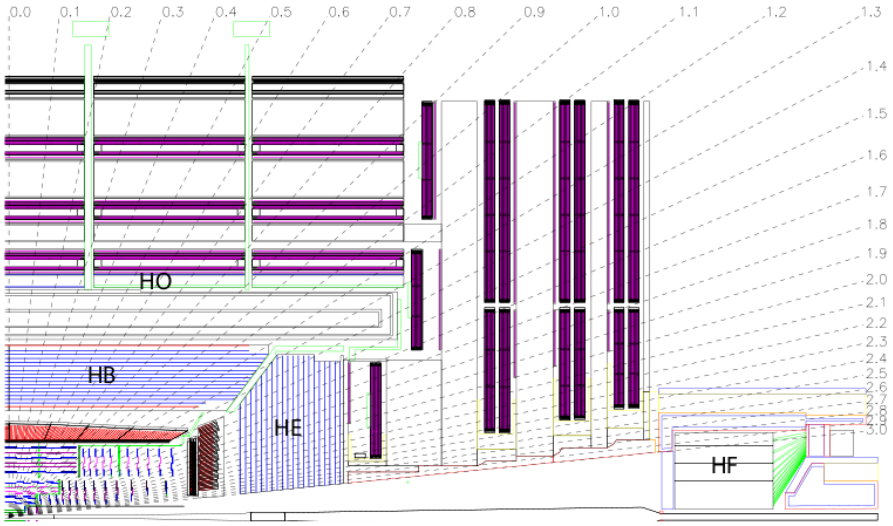


Fig. 15. Sketch of the longitudinal view of the CMS detector, showing the main components of the hadronic calorimeter.

$H \rightarrow ZZ \rightarrow llll$ where l is a lepton. If the leptons are muons, we will obtain the best accuracy for the Higgs mass reconstruction due to the low radiation losses of muons when they interact with the detector material, compared to say, electrons.

The system for muon detection is composed of three different types of gaseous detectors: drift tube (DT) chambers, cathode strip chambers (CSC) and resistive plate chambers (RPC). The muon system is interleaved with the iron return yoke for the magnetic field. The objective of the muon system is to identify muons of low and high energies and measure their momentum. The different radiation environments and the large detector area to be covered have been the main reasons to choose different detector technologies for the muon system.

V.6.1 Drift Tubes Chambers

A DT chamber is a device used to measure the spacial position of charged particles. The basic DT unit consists of a cell filled with an ionizing gas. At CMS the DT cells have an anode wire running in the middle, cathode strips on the sides and electrode strips at the top and bottom of the cell as shown in Figure 16. When a charged particles passes through the DT cell, it ionizes the gas leaving a trail of electrons and ions that produce a signal collected by the anode.

The barrel muon section of the CMS detector is composed of five wheels, each of them divided in 12 identical sectors. The DT chambers cover the barrel region ($\eta < 1.2$) where the muon rate is low and the residual magnetic field in the chambers is low. They are organized in four layers (MB1, MB2, MB3 and MB4) per wheel as is shown in Figure 17. The three inner layers (MB1, MB2 and MB3) host 12 chambers each. The MB4 layers host 2 chambers at the top and bottom of each wheel, giving a total of 14 chambers for this outer layer.

V.6.2 Cathode Strip Chambers

The CSC system is composed of 468 units equally distributed in two endcaps. Each CSC has a trapezoidal shape and is symmetrically distributed in concentric

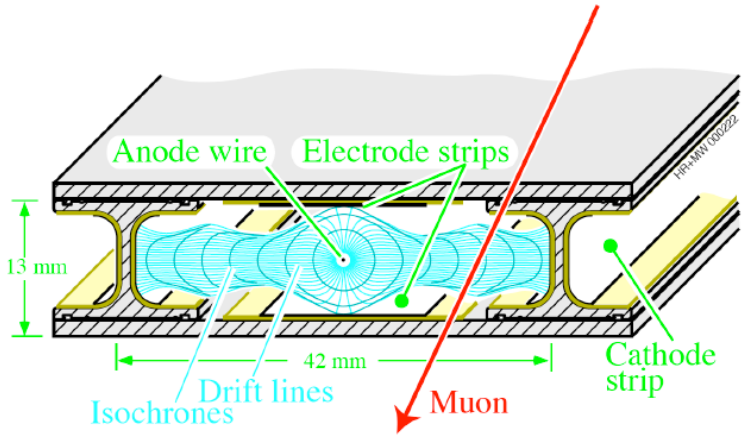


Fig. 16. Sketch of the basic unit cell of a DT chamber at CMS. A voltage of +3600 V is a applied to the anode wire, 1800 V for the electrode strips at the top and bottom of the cell and −1200 V for the cathode strips on the sides [11].

rings around the beam line (Z coordinate). The Chambers consist of six detecting layers known as gas gaps. Each gas gap has a layer of cathode strips radially distributed. A set of anode wires runs almost perpendicular to the cathode strips as shown in Figure 18. When charged particles pass through the chambers, the gas ionization generates an avalanche of electrons which produce a charge on the anode wire and a subsequent image charge on the a group of cathode strips [11]. Because the signal of the anode wires is very fast, it is used for the level one trigger (see Section V.6.4). CSC chambers were used for the two endcaps ($\eta < 2.4$) where the muon rate, the background and the magnetic field are high.

V.6.3 Resistive Plate Chambers

In CMS the basic module of an RPC chamber consist of two parallel plates polarized with high voltage and filled with an ionizing gas. The readout of the signals left by charged particles passing through the chamber is performed with a set of metallic strips as shown in Figure 19.

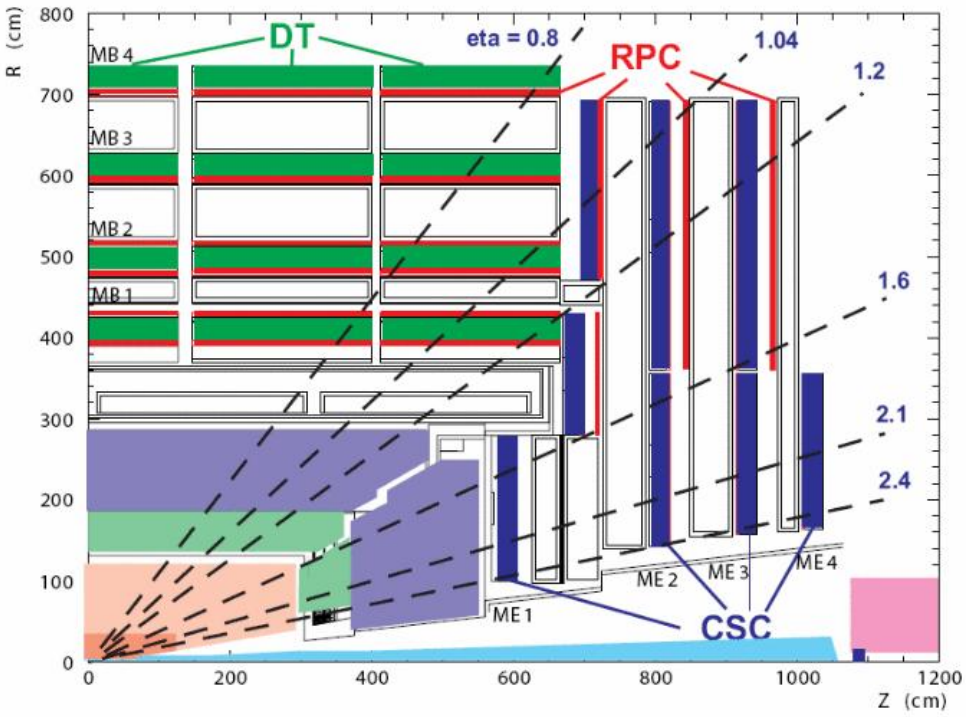


Fig. 17. Layout of one quarter of the CMS muon system. The muon detection system of CMS is composed of three different types of detector: drift tube (DT) chambers, cathode strip chambers (CSC) and resistive plate chambers (RPC). The different technologies are used in order to have a good momentum and position resolution and a fast response for the trigger system (see Section V.6.4). Also, the different radiation environments and the large detector area to be covered are some of the main reasons to choose different detector technologies for the muon system.

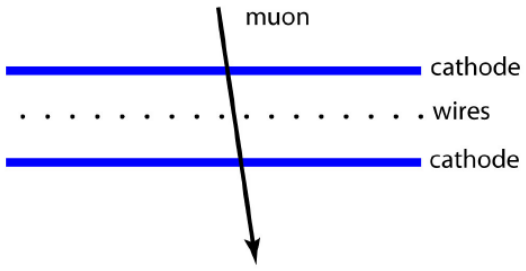


Fig. 18. Sketch of a single gas gad of a CSC chamber [11].

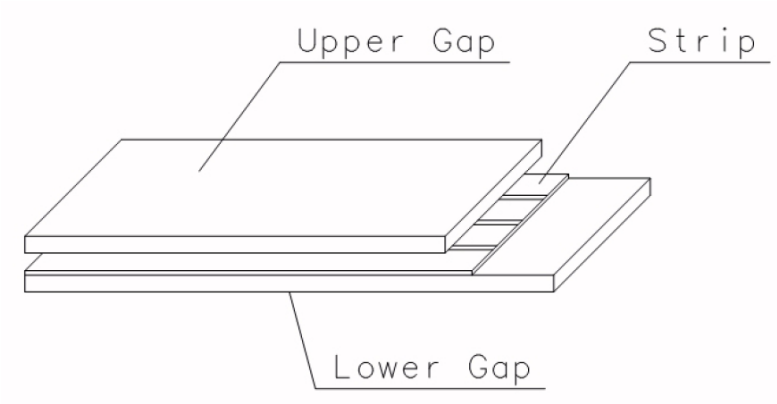


Fig. 19. Sketch of a the basic module for an RPC chamber [11].

The RPC chambers provide a good spatial and timing resolution for charged particle detection. The RPC chambers are located in both detector regions: barrel and endcaps. Due to their fast response to radiation, the RPC chambers were added to the muon system to provide additional trigger information to complement the information coming from the DT and CSC chambers.

For the Barrel section, a total of six layers of RPC are installed. As is shown in Figure 17, there are two layers of RPCs at stations MB1 and MB2 and just one layer for M3 and M4. On each endcap there are four layers of RPC, one per each CSC. The RPC chambers cover a region of $\eta < 1.6$.

V.6.4 Trigger and Data Acquisition System

The collection of data and preliminary identification of events of interest is carried out using a Trigger and Data Acquisition (DAQ) system. In CMS, at the nominal luminosity, the expected bunch crossing frequency is 40 MHz with an average of 17 proton-proton interactions per crossing. A first level trigger (L1) makes an initial filter for interesting events, searching for interactions of muons, electrons, photons and jets that pass an initial threshold of transverse (to the beam) energy and momentum. The L1 looks as well for missing transverse energy that may indicate interesting physics (see Section VI.1). A second trigger level, called the high level trigger (HLT),

is applied to the remaining events after the first level trigger is applied. This high level trigger performs a more detailed selection of events by applying a stricter criteria. In order to minimize the time to discard events without interesting information, the HLT has been divided into a set of virtual trigger levels. Initially, information is used from the muon and calorimeter detectors, followed by the inclusion of the pixel detector, and finally all of the tracker information.

At nominal design values, the L1 trigger will operate at a frequency rate of 100 KHz. The HLT will input the 100 kHz coming from the L1 and reduce it to 100 Hz for the final data storage. During the 2010 data taking by CMS, the L1 trigger operated at a rate of 70 KHz and the HLT achieved output rates between 300 and 600 Hz.

The DAQ system has been designed to operate in a synchronous way with the trigger in order to sustain the input rates coming from the L1 and provide the necessary computing power for the HLT. Each CMS sub-detector has an electronic system, known as the Front Electronic System (FES), to store the data in 40 MHz pipelined buffers [11]. The data are extracted from the buffers by the L1 trigger in a synchronous way using a Timing, Trigger and Control (TTC) system. The TTC system synchronizes the detector readout with the L1 trigger rate. The data of each sub-detector are sent from the FES buffers to the DAQ system using a set of electronic cards known as Front-End-Drivers (FED's) where the information is digitized and organized. The FED's put the information in data words of 64 bits using a format that starts with an 8 byte word known as "header" that marks the beginning of an event, followed by a set of 8 byte words where the relevant information of the detector is encoded, and finishing with an 8 byte word known as "trailer" that marks the end of the event. Each FED is connected to an S-Link64 card where the data words are checked for transmission errors (this is known as a Cyclic Redundancy Check (CRC)). From the S-Link64 cards, the data words are transferred to a Front-end Read-out Link (FRL) which provides an optical link used to send the information to the computer farms where all the pieces of information of one event are assembled.

A CRC check is also performed by the FRL before sending the information to the event builder system in order to detect transmission errors for the S-Link64 card.

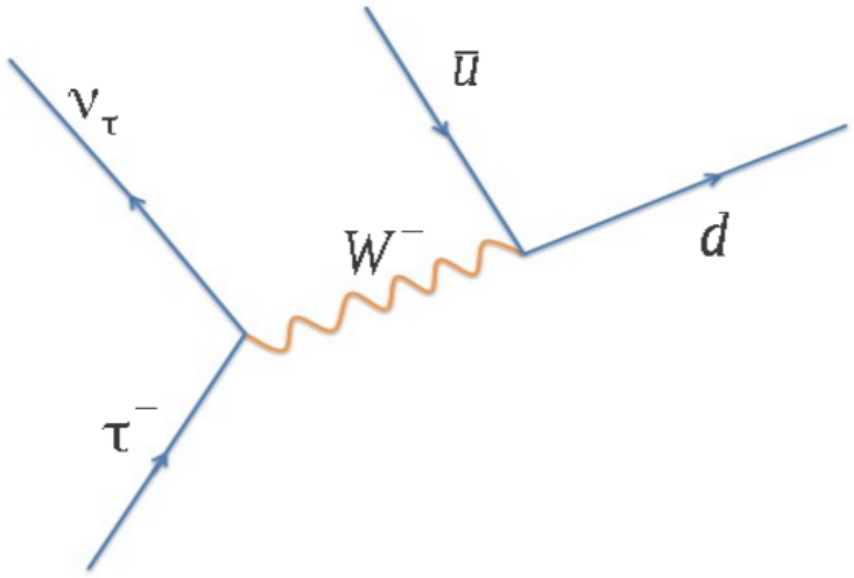


Fig. 20. Leading order Feynman diagram of the hadronic decay of a tau to a charged pion and a tau neutrino.

V.7 Introduction to Tau Physics

Taus are very interesting and challenging particles to study. Due to their heavy mass, taus are the only particles of the lepton family that can decay hadronically and leptonically [12]. A tau decays to a tau neutrino through the emission of a W boson, that can decay hadronically or leptonically. When the tau decays hadronically, the W boson decays to a quark and anti-quark pair, $(u, \overline{d})$, $(\overline{u}, d)$ or $(c, \overline{s})$ depending on the charge of the W, which will result in a jet of one, three or five charged hadrons plus neutrals depending on the fragmentation process of the quarks. A similar process happens for leptonic decays of the tau, but in this case the W boson decays to an electron or a muon and associated neutrinos.

The branching ratio of taus decaying to a jet of charged hadrons plus neutrals is around 64%. Out of the 64%, approximatelly 84% corresponds to decays to one charged hadron and neutrals. Leptonic decays to muons or electrons have roughly the same branching ratio, around 17.5%. Figures 20 and 21 show the Feynman diagrams for a hadronic and leptonic decay of the tau.

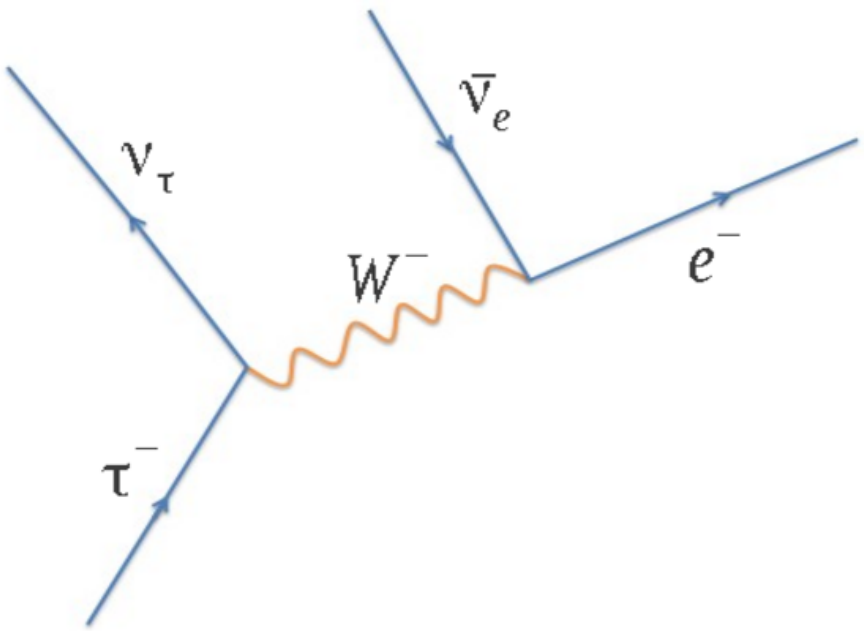


Fig. 21. Feynman diagram of the leptonic decay of a tau to an electron and electron anti-neutrino and a tau neutrino.

Taus are heavy particles and have a short life time of $2.9 \times 10^{-13}s$. For this reason their decay length is around 87 μm which means that they can not reach the first layer of the pixel detector. This implies that the detection of taus is done indirectly and a good understanding of the tau signatures and decay topologies is needed in order to distinguish them from other particles with similar characteristics. In Section VII.3 we will describe the tau reconstruction and identification criteria that we use for our analysis.

V.7.1 The Importance of Tau Physics

Several physics models beyond the standard model favor the discovery of new physics in final states containing tau leptons [2]. Some new resonances with standard model couplings (*i.e.* Z') are most likely to be found in decays to electrons or muons, but the decay to taus is still important in order to test the universality of the couplings [3].

For a Higgs boson with standard model couplings, the decay to taus for a mass under 200 GeV has the second highest decay fraction (branching ratio) among all possible decays as shown in Figure 22. Even though the decay to events coming from the $b\bar{b}$ quarks has a higher branching ratio, it is more difficult to make a discovery in this channel because the rate of events that mimic this type of decay is very high. Therefore taus are a better candidate to discover the Higgs in this low mass regime as fake rates are lower than for processes coming from $b\bar{b}$.

Another interesting topic, where taus can play an important role in the discovery of new physics, is dark matter detection at the LHC. Currently, we know that barionic matter, composed of protons, neutrons and electrons, corresponds to just 4% of the total estimated matter of the universe. The remaining fraction is divided into dark matter (around 23%) and dark energy. The existance of dark matter and its effects have been studied indirectly by astrophysics through its effects on visible matter and the expansion of the universe.

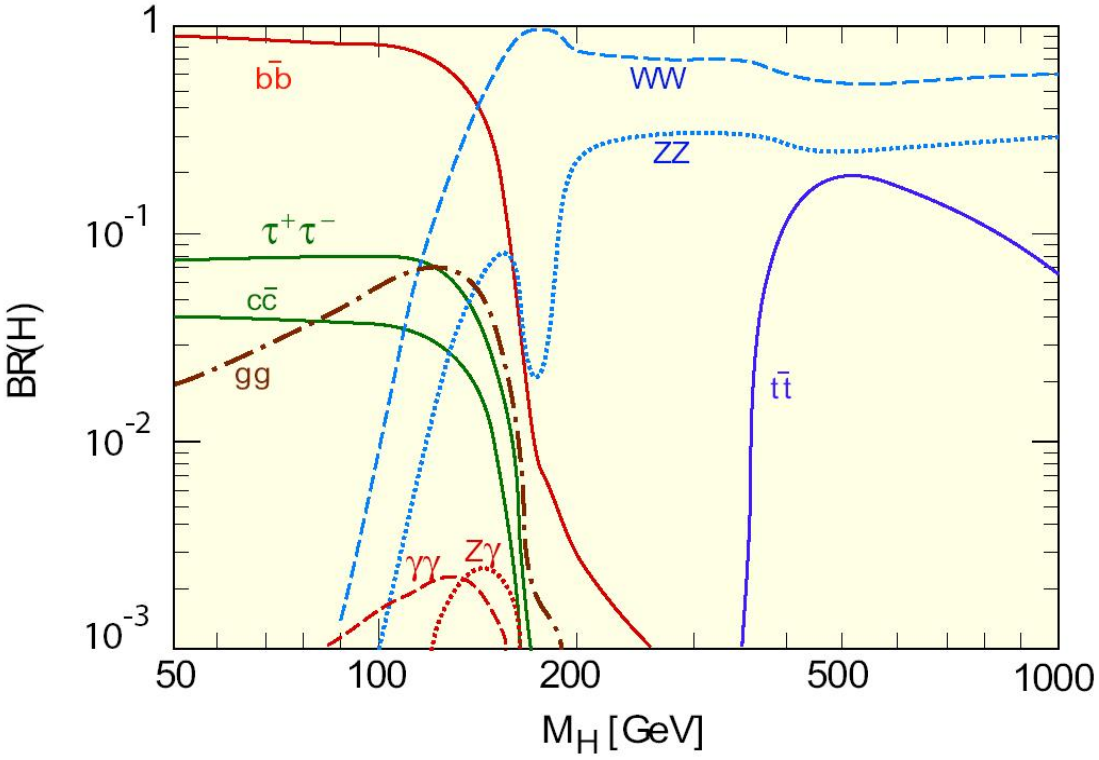


Fig. 22. Branching ratio of Higgs decays versus different possible Higgs masses. The decay for taus has the second highest branching ratio for a mass under 200 GeV. Although the decay to $b\bar{b}$ has the highest probability to occur, there are larger backgrounds than taus. For this reason, taus present one of the best opportunities for an eventual discovery of the Higgs boson with a mass under 200 GeV.

Supersymmetry (SUSY) is an extension of the SM that aims to answer some of the open questions in the SM. Some models based on SUSY give rise to a neutral particle called the neutralino one ($\tilde{\chi}_1^0$), which is the lightest supersymmetric particle. It has been proposed that this particle is the fundamental component of dark mater. In SUSY, there is a free parameter called $tan\beta$ which is defined as the ratio of the vacuum expectation value of two neutral Higgs fields. In the SUSY model, the Higgs field is composed of five particles, three with an electric charge and two neutral. According to SUSY, if $tan\beta$ has a large value the production mechanism that gives rise to the neutralino makes it most likely to be detected in final states containing taus (with a probability close to 100 %) [13]. Nevertheless, if SUSY is found in other final states at low $tan\beta$, the decay to taus is still important in order to test the universality of the couplings.

CHAPTER VI

ANALYSIS DESCRIPTION AND STRATEGY

We search for any new heavy resonance decaying to a pair of taus. In order to have a reference point for our analysis, we choose a Z' gauge boson as a benchmark. We followed the model where this boson is produced in a quark and anti-quark annihilation process using the same physical characteristics as the electroweak Z boson but with a larger mass.

Since taus can decay hadronically or leptonically, there are several possibilities to search for this resonance. The search can be performed looking for one tau decaying to a muon and and the other to a jet of hadrons and associated neutrinos. Furthermore, we can search for one tau decaying to an electron and the other to a jet of hadrons and associated neutrinos or both taus simultaneously decaying hadronically or leptonically. For our search, we used the tau decay mode where one tau decays to a jet of hadrons and a tau neutrino, and the other decays to an electron and associated neutrinos. We will refer to this as the $e\tau$ channel ($Z' \rightarrow \tau^+\tau^- \rightarrow e + \tau_h + X$).

The analyses of taus with decay modes other than the $e\tau$ channel were carried out by different groups from several Universities. To maximize our statistical power, the analyses were combined utilizing a methodology that is explained in detail in Chapter X. Final states such as $Z' \rightarrow \tau\tau \rightarrow ee/\mu\mu$ where the two taus decay leptonically to the same type of leptons are not considered due to the difficulties involved in distinguishing them from direct production of di-electrons or di-muons, and the copious $Z \rightarrow ee/\mu\mu$.

VI.1 Missing Transverse Energy

The total momentum for the proton-proton collision in the center of mass frame is nearly zero. For this reason, the total momentum of all particles produced after the collision should be close to zero as well. Neutrinos cannot practically be measured by our detectors because they are neutral particles that just interact weakly. In order

to compensate for this effect we use a physical quantity known as missing energy in order to perform an indirect measurement of the momentum imbalance due to the undetected neutrinos and other neutral particles (e.g neutralinos).

The total momentum contains all the particles that can be detected (visible) and the ones that do not leave a signal in our detector (invisible). By the law of momentum conservation, the total momentum of all particles produced after the collision should be equal to the initial momentum at the center of mass frame as shown in Equation 34.

$$\sum \overrightarrow{p}^{total} = \sum \overrightarrow{p}^{visible} + \sum \overrightarrow{p}^{invisible} = \overrightarrow{0} \tag{34}$$

The missing energy is then defined as the negative sum of the momentum of the detected particles as shown in Equation 35.

$$\overrightarrow{E}_{miss} = -\sum \overrightarrow{p}^{detected} \tag{35}$$

Because the proton-proton collisions happen along the Z direction, many particles are detected in the forward region of the detector as result of soft scattering processes. Events with a high component of momentum in the $X-Y$ plane (transverse region) and low momentum in the longitudinal direction are of great interest as those normally come as result of hard scattering processes. For this reason we look for events with a high component of transverse momentum and transverse missing energy, one on the main signatures of the Z' decay topology. The missing transverse energy is defined as shown in Equation 36.

$$\overrightarrow{E}_{miss\ T} = \overrightarrow{\not{E}_{\mathrm{T}}} = -\sum \overrightarrow{p}_{T}^{detected} \tag{36}$$

The mismeasurement of the missing transverse energy can lead to an imprecise identification of physics processes of interest. Effects related with physical processes, the detectors performance and the acceptance region it covers, can lead to errors in the estimation of the missing transverse energy [14]. For example, cracks in the detector acceptance region, noisy towers in the calorimeters or bad modules

in any sub-detector in general can be a source of error. A good understanding of physical processes such as pile-up (PU) and underlying-events (UE) is essential for a correct estimation of the missing transverse energy as well. By pile-up we refer to the other proton-proton interactions that occur in addition to the central collision. Underlying events correspond to particles that result from final or initial state radiation processes as described in Section VI.2. The correct separation of particles coming from pile-up and underlying-events from the particles produced in the hard collision is fundamental for an accurate estimation of the missing transverse energy.

VI.2 Backgrounds

Physics events with similar signatures and decay topologies that fake the signal of interest are known as backgrounds. By signal we refer to the Z' boson in the tau decay mode explained above. It is essential to quantify accurately the expected number of background events in the mass region where we expect to observe our signal.

In order to suppress backgrounds, we apply a set of selection criteria to discriminate them from our signal. These selections should be robust to reduce the backgrounds as much as possible without removing a significant fraction of the expected events from our signal. Also, we chose selections that do not introduce large systematic effects and that are not correlated with other selections if possible. All the selection criteria for this analysis are explained in Section VII.6.

We considered the following backgrounds:

- $Z \rightarrow \tau\tau$: This background has the same signatures as the $Z' \rightarrow \tau\tau$, nevertheless the Z is easy to discriminate as it sits in a low mass region ($M_{Z^0} = 91.19 \pm 0.0021$) compared to the Z' which is expected to have a mass above 400 GeV.

  This process is also used as a control region in order to validate our tau identification criteria (see Section VII.3). The physical properties of the Z boson have

been accurately measured in previous experiments and those measurements have been successfully reproduced by CMS [15]. For this reason our selection criteria have been chosen such that by relaxing a few cuts we can observe the $Z \rightarrow \tau\tau$ mass peak.

- $Z \rightarrow ee$: This background enters into our signal region when a real electron from the Z passes all the electron identification criteria and the hadronic tau is faked by a badly reconstructed electron.

- $t\bar{t}$: The top (t) and anti-top ($\bar{t}$) quarks decay primarily to a bottom quark (b) or an anti-bottom ($\bar{b}$) respectively through the emission of a W boson. This process produces signatures similar to the $Z' \rightarrow \tau\tau \rightarrow e\tau$ decay: the hadronic W boson decay can look like a tau jet and the semielectronic decay can look like the semielectronic decay of a tau. Also, the W boson can decay to a tau and a tau neutrino that can mimic the leptonic or hadronic tau from the Z' decay. The associated jets from the b or $\bar{b}$ quarks hadronization though tend to produce a lot of extra particles that are used to identify this background. The b-tagging algorithm used to eliminate $t\bar{t}$ events is described in Section VII.6.

- $W + jets$: These events enter into our selection when a W and an uncorrelated jet coming from a different process fake the leptonic (electron) or hadronic tau. The jet can come from different sources such as Initial State Radiation (ISR), Final State Radiation (FSR) or QCD processes. The ISR radiation happens as result of the interaction of quarks before the collision of the proton beams through the emission of gluons or photons. Similarly the FSR results as consequence of radiations of gluons or photons from decay products of particles produced after the pp collision. Figure 23 shows the Feynman diagrams for ISR and FSR processes.

- $\gamma + Jets$: This background enters into our selection when a high energy photon (γ) fakes the leptonic tau and a jet fakes hadronic tau. The photon and the jet can come from different sources such as ISR, FSR or electromagnetic

interactions among quarks. These events do not have intrinsic missing energy and due to conservation of momentum the photon and the jet are expected to be back to back.

- *QCD*: In these events, electrons come from heavy flavor decays ($b\bar{b}$ || $c\bar{c} \rightarrow e + \nu_e + X$) and conversions in the material coming from photons associated with jets. The hadronic tau is faked by jets coming from the fragmentation of quarks and gluons. These jets can also produce heavy hadrons that decay to real taus (*i.e.* $D^+ \rightarrow \tau^+\nu_\tau$, $D^+ \rightarrow \pi^+\pi^0$, $D^+ \rightarrow k^+\pi^0$, $B^+ \rightarrow \tau^+\nu_\tau$).

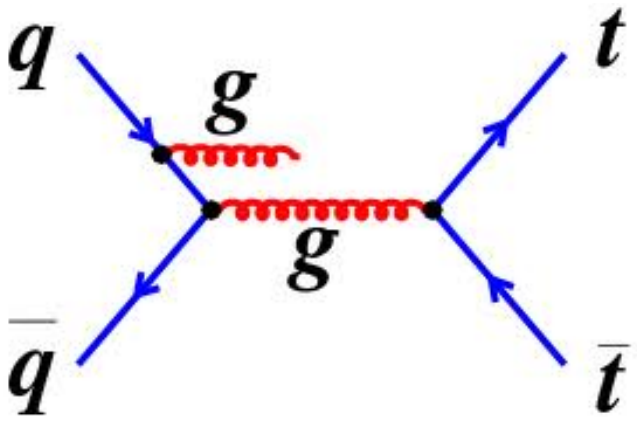
In Figures 24 to 29 we show the Feynman diagrams that represent the production mechanisms used to simulate the backgrounds described here.

VI.3 Corrections in the Simulation for Pile-up Effects

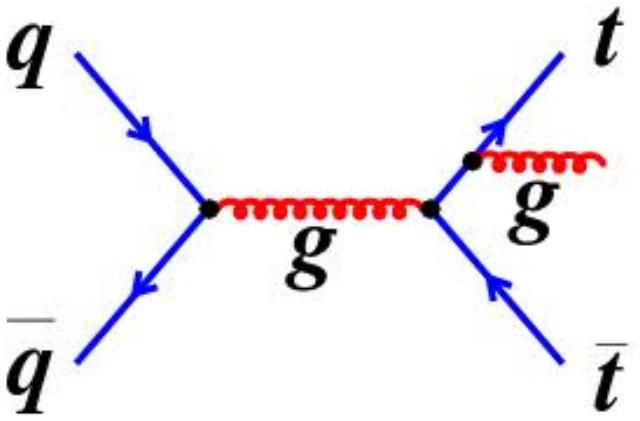
The simulated signal (Z') and the background samples described in Section VII.2 do not include pile-up (PU) effects. To compensate for the differences between the collision data and the simulation in the missing transverse energy distributions due to pile-up effects, we introduced a correction factor as shown in Equation 37.

$$\overrightarrow{\not{E}_{\mathrm{T}}}^{Corrected} = \overrightarrow{\not{E}_{\mathrm{T}}}^{Raw} + \overrightarrow{\Delta\not{E}_{\mathrm{T}}}^{PU} \tag{37}$$

In order to quantify the correction factor for PU, gamma+jets events are analyzed [17]. Because gamma+jets events do not contain any real intrinsic missing energy, observation of non zero missing transverse energy values is entirely due to mismeasurement of jets, mismeasurement of photons and PU. In a perfect world, gamma+jets events are back to back. Nevertheless, the mismeasurements in jets, photons and PU effects create a topology where gamma+jets events are not back to back. By projecting the momentum of the jet onto the direction of the photon, we can measure the effects due to PU and the mismeasurement in the jet energy. The perpendicular component of the jet projection represents the effects in the jet direction due to PU and the longitudinal component quantifies the mismeasurement

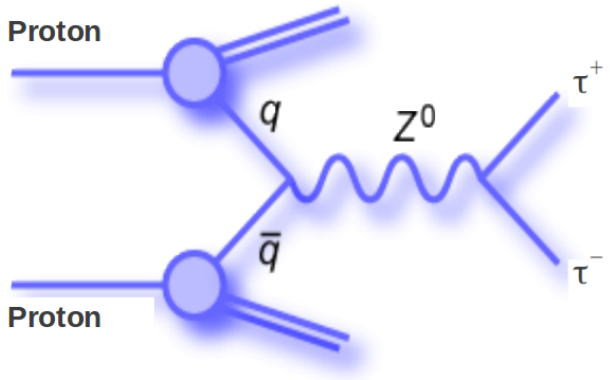


(a)



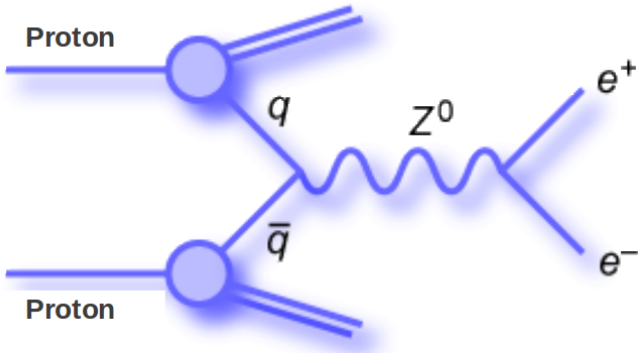
(b)

Fig. 23. Feynman diagrams for initial state radiation (a) and final state radiation (b).



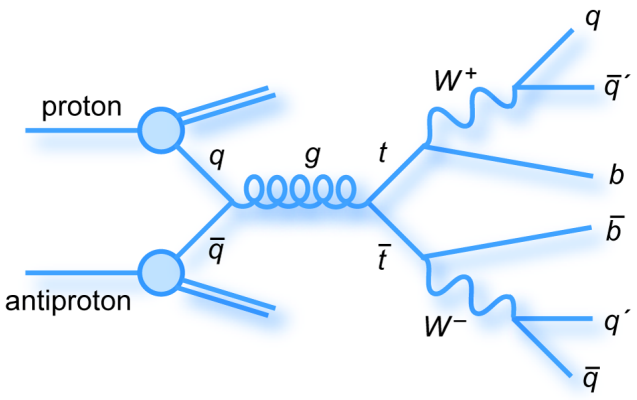
(a)

Fig. 24. Feynman diagram for the $Z \rightarrow \tau\tau$ background [16]. This background is also used as a control region in order to validate the selection criteria used for tau identification. Notice that this background has the same topology as the Z' but we can easily discriminate it because the mass of Z boson a small compared with the expected mass of the Z'.



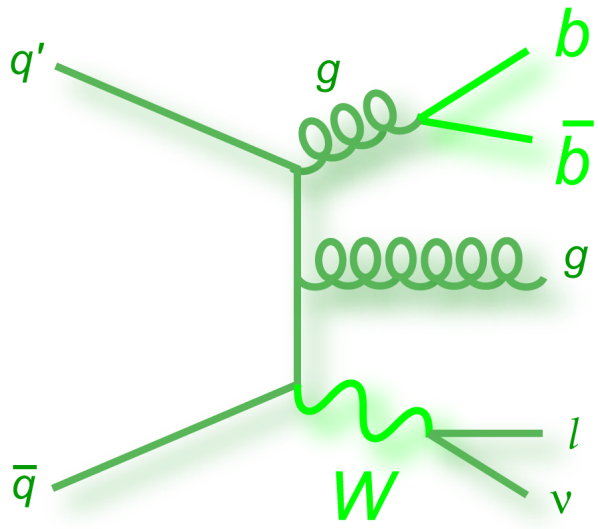
(a)

Fig. 25. Feynman diagram for the $Z \rightarrow ee$ background [16]. This background enters into our selections because we have a real electron from the Z faking the electron coming from the leptonic tau and a badly reconstructed electron from the Z faking the hadronic tau.



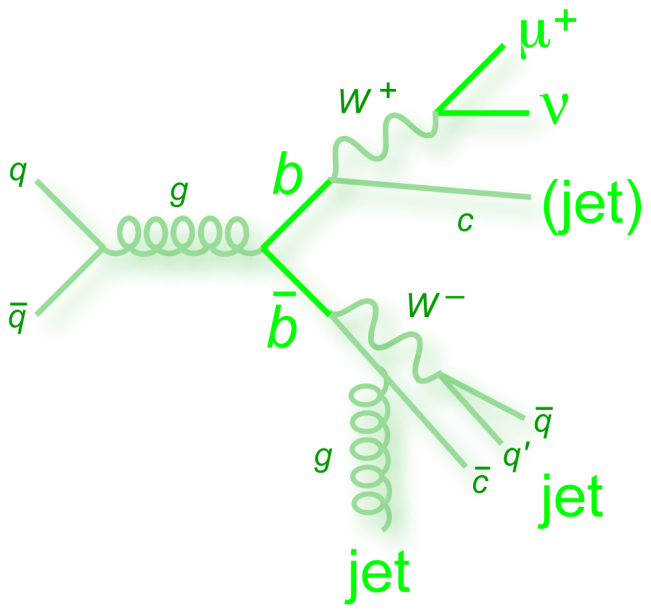
(a)

Fig. 26. Feynman diagram for the $t\bar{t}$ background [16]. This background enters into our selection when a clean electron from the W or a jet from the b quark fakes the electron from the leptonic tau and a jet from b quark fakes the hadronic tau.



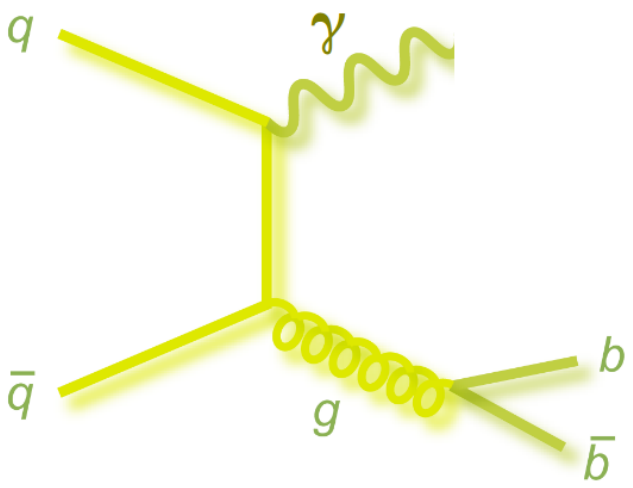
(a)

Fig. 27. Feynman diagram for the $W + Jets$ background [16]. This background enters into our selection when a W decays leptonically producing a real electron or decays hadronically faking the tau jet and an uncorrelated jet from initial or final state radiation process fakes the hadronic tau or the electron.



(a)

Fig. 28. Feynman diagram for the QCD background [16]. This background enters into our selection when the hadronic and leptonic taus are faked by jets produced in hard scattering.



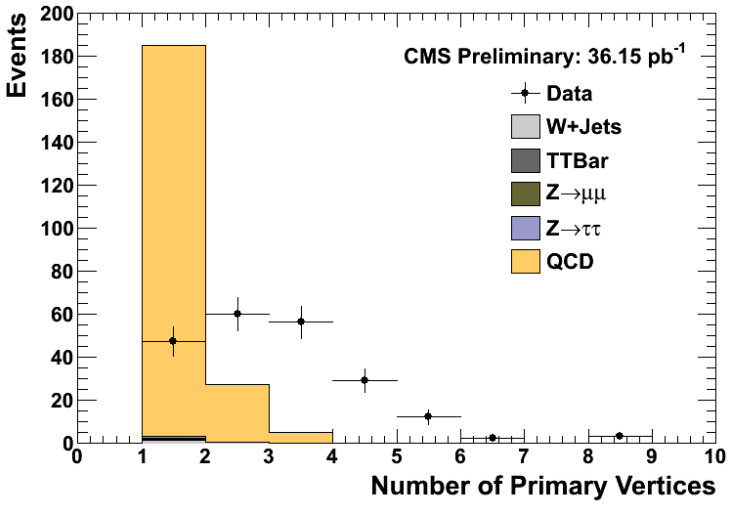
(a)

Fig. 29. Feynman diagram for the $\gamma + Jets$ background [16]. This background enters into our selection when a high energy photon (γ) fakes the leptonic tau and a jet fakes hadronic tau. The photon and the jet can come from different sources such as ISR, FSR or electromagnetic interactions among quarks.

in the jet energy scale. The resolution loss is scaled (see Equation 38) as a function of the number of reconstructed primary vertices in an event.

$$\delta \not\!\!E_{\mathrm{T}} = \sqrt{n} \cdot \sigma_{PU} \cdot F_{scale}(\not\!\!E_{\mathrm{T}}). \tag{38}$$

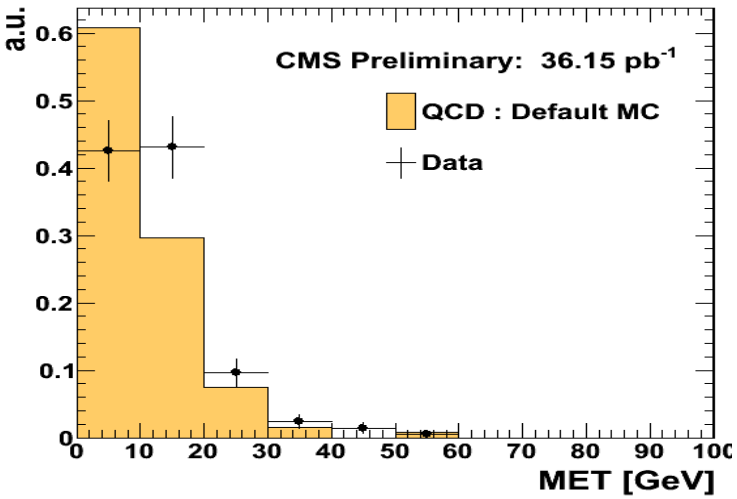
In Equation 38, σ_{PU} represents the pile-up distribution, n the number of vertices, and $F_{scale}(\not\!\!E_{\mathrm{T}})$ represents the jet energy scale factors. The distribution for the number of vertices n is obtained directly from data using background enhanced regions as shown in Figure 30. The technical implementation of the PU correction is done by adding energy to the MC sample on an event by event basis using Equation 38 and the probability density function for the number of vertices.



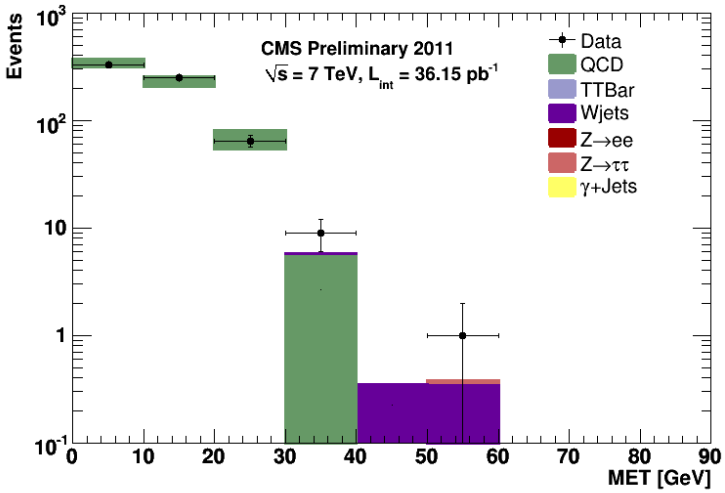
(a)

Fig. 30. Distribution of the number of vertices in data, compared with the distribution for an enhanced region with QCD events without introducing any PU corrections.

Figure 31 shows the missing transverse energy distribution before and after introducing the corrections for pile-up. Notice the good agreement in the distribution after the correction. The selections used to obtain a region dominated by QCD events is described in Chapter VIII.



(a)



(b)

Fig. 31. Missing transverse energy before and after corrections to include PU effects. (a) distribution before applying the correction. The distribution is for the final state where one tau decays to a muon and associated neutrinos and the other to a jet of hadrons ($\mu\tau$ channel). (b) distribution after corrections for the $e\tau$ channel. Notice the good agreement between data and MC after applying the correction for PU.

VI.4 Mass Reconstruction

We can not perform the full mass reconstruction of the Z' boson due to our imprecise knowledge of the energy of the missing neutrinos. In order to compensate for the neutrinos, we incorporate the missing transverse energy into our calculation of the invariant mass from the visible decay products:

$$M(e, \tau, \not{E}_{\mathrm{T}}) = \sqrt{(E^e + E^\tau + \not{E}_{\mathrm{T}})^2 - (\overrightarrow{p}^e + \overrightarrow{p}^\tau + \overrightarrow{\not{E}}_{\mathrm{T}})^2} \tag{39}$$

Including the missing transverse energy improves the resolution of the invariant mass estimate due to the intrinsic missing energy in the Z' decay topology and therefore the discriminatory power against backgrounds with a low or non missing energy component.

Figure 32 shows the mass distributions using simulated data for a Z' with a mass of 500 GeV and all the associated backgrounds. The plot is made after applying a set of selection criteria that will be outlined in Section VII.6 and uses the reconstruction formula described above. Notice that most of the backgrounds are located in the low mass region of the mass spectrum. In the high mass region the expected signal is fairly clean with some small contamination coming from $t\bar{t}$ and $W + Jets$ events.

VI.5 Strategy

The Z' boson is a neutral resonance expected to have a large mass. For this reason, we look for a pair of taus with high transverse momentum, opposite electric charge and decay products pointing nearly in opposite directions within the fiducial volume of the detector acceptance. The associated neutrinos in the tau decay modes we are studying can create missing transverse energy in the expected events from the Z' boson. The analysis has been designed to select events that meet the characteristics described above, minimizing the presence of backgrounds in the regions where a possible signal can be detected. Our selection criteria to target these events have been chosen such that correlations and systematic uncertainties are minimized.

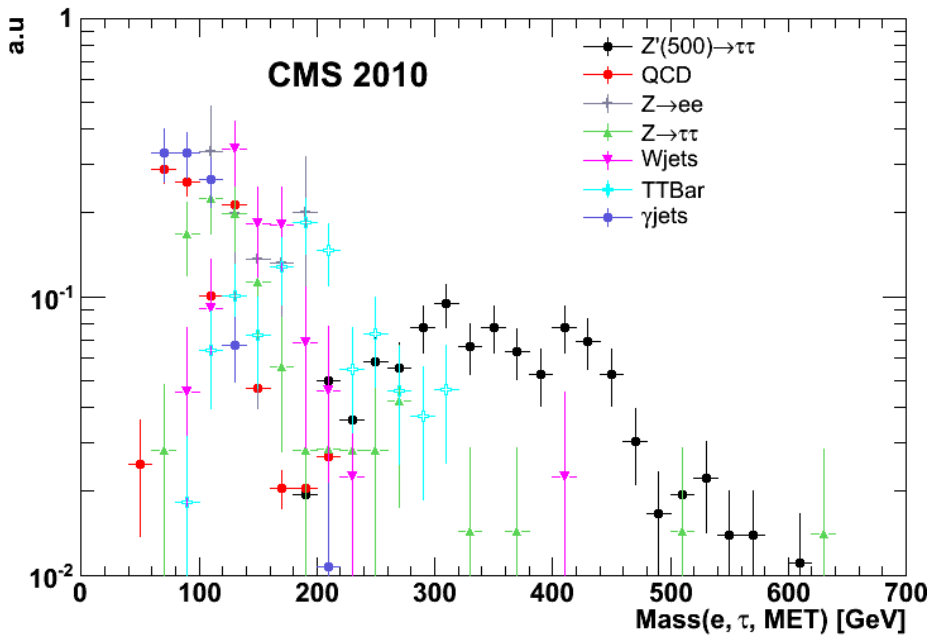


Fig. 32. Mass reconstruction using simulated events for backgrounds and signal. In this figure we used as signal a Z' particle with an expected mass of 500 GeV. The distribution are after all the selection criteria outlined in Section VII.6 have been applied.

CHAPTER VII

SIMULATION, RECONSTRUCTION AND SELECTION CRITERIA

In order to find new physics, we need to carefully design the selections that will allow us to distinguish the characteristics of the events of interest (also referred as signal). The design process of the physical selections is carried out using simulated (Monte Carlo) data containing the final states of the particles we are interested in. Another important feature is the inclusion of simulated events with known physics processes that can potentially pass our selection criteria and fake our signal.

VII.1 Simulation of Physics Events and Detector Effects

In general a simulation can be defined as a process to model a specific situation of interest. In particle physics we use Monte Carlo techniques in order to describe physical and mathematical processes. The modeling of these processes can be understood through previous experimentation or can come from the description of models proposing new physics. In either case, we can not predict with 100% certainty the outcome of an experiment, all we can do is to assign a probability for an event to occur. For this reason a Monte Carlo simulation is used to solve complex numerical problems where exact solutions are not possible.

For example, although we understand physics processes such as $Z \rightarrow \mu\mu$ or $Z \rightarrow \tau\tau$, we can not predict exactly their production in a proton-proton collision. Therefore, we assign a probability for these events to occur under specific physics conditions. The Monte Carlo simulation is used to compare the experimental results with the data and to spot any possible problems in the detector performance and the selection criteria used to obtain or model the physics events of interest.

It is important to model correctly known physics processes that have been theoretically explained and experimentally understood. Processes such as particle scattering, development of an electromagnetic or hadronic shower, motion of relativistic particles in a magnetic field, interactions among particles, fragmentation of quarks

(very difficult to model) and detector effects, are essential to test different aspects of an experiment in particle physics. Furthermore, it is important to properly model the production mechanisms through which different particles are produced such as quark and anti-quark annihilation, strong interactions among quarks, electromagnetic radiation etc.

The modeling of possible new physical processes is also fundamental in particle physics. This requires a good understanding of the possible production mechanisms of new physics, a correct description of the properties of the new particles and how they can eventually be detected. For example, the simulation of the Z' boson in this thesis requires the particle to be produced in a quark and anti-quark annihilation process with standard couplings and to have a high mass. Essentially, the simulation of the Z' has the same characteristics as the known electroweak Z boson but with a larger mass. The possible tau decays were introduced using the correct branching ratios for the different tau decay modes.

To perform the complete simulation of physics events, CMS proceeds in two different steps. The first step corresponds to the generation of particles from specific physics events resulting as consequence of the proton-proton collisions. In this process, particles are produced with the correct momentum, energy and position depending on the initial conditions of the collision. The second step consists of introducing detector effects that occur as a consequence of the different interactions of the particles with the different components of the CMS detector. These interactions will change the initial values of position, energy and momentum of the particles introduced at the generator level. The precise understanding of the detector response is important for the reconstruction of real particles produced in the proton-proton collision, in order to apply corrections to the measured momentum, energy and position.

In addition to detector effects, the inclusion of pile-up and underlying events in the simulation is important to emulate correctly the physics processes involved in the collision and detection of real particles.

The proton-proton collisions at the LHC involve very complex processes between quarks and gluons inside the protons (partons). Although protons are formed by two "up" quarks and one down quark "d", there are many underlying processes in the internal structure of the proton, such as a constant exchange of gluons that hold the quarks together, that are very difficult to model. Therefore, the total energy of a proton will be split among the different partons and gluons. The exact distribution of the energy among the participants is not exactly known. For that reason, the energy distribution is modeled using Particle Density Functions (PDF's), that attempt to describe, on average, the distribution of energy depending on the initial physical conditions of the protons.

The most simple and probable interactions among partons that yield the production of specific particles is known as a Leading Order process (LO). More complex interactions, where for example the quarks have final or initial state radiation processes, are known as Next to Leading Order process (NLO). The NLO order process is more difficult to model and calculate mathematically [18]. There are Next to Next to Leading Order processes (NNLO) that become cumbersome to simulate and calculate.

The algorithms used to identify particles in real data are tuned using a robust simulation to the important physics processes described above. In order to determine the effectiveness of the algorithms used to identify specific particles, the initial parameters of the generated particles are compared with the reconstructed objects in the simulation. A reconstructed object at the simulation level can be defined as a particle candidate pre-identified using specific algorithms, after detector and other physics effects have been introduced. The efficiency is then defined as the ratio between the reconstructed particles in the simulation that are matched to a generator level particle which passed the algorithm, with respect to all the matched reconstructed particles. The matching is often performed by selecting the closest reconstructed candidate of interest to a generator level particle using a cone in $\eta - \phi$ space ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$).

The majority of Monte Carlo samples in our analysis were produced with the PYTHIA [20] event generator. The PYTHIA generator can simulate parton processes at leading order level, including the fragmentation of quarks and gluons (hadronization), that give rise to production of various particles.

The MADGRAPH [21] event generator was used in one of the Monte Carlo samples ($W + Jets$ see Section VII.2). This event generator allows the introduction of additional hard parton radiation process.

The TAUOLA [22] package was used to introduce the different tau decay modes, with correct production probabilities. This package also contains the polarization of taus which is needed in order to obtain the correct momentum distribution of these leptons decay products.

Finally, the GEANT4 [23] simulation package was used in order to introduce all the detector effects due to the passage of elementary particles through matter.

VII.2 Datasets and Monte Carlo Samples

The data used for this analysis is from the November 4^{th} CMS official validation, generated and reconstructed using the CMS Software (CMSSW) framework CMSSW_3_8_6. Only data from "good" run ranges and lumi sections approved by the CMS Data Quality Monitoring group (DQM) were used.

Official Monte Carlo samples from the Fall 2010 production were used. The PYTHIA event generator was used for the the Z^0, $t\bar{t}$, and QCD (electromagnetic enriched) production, whereas MADGRAPH was used for $W + Jets$ processes.

The data sets used for collision data and simulations, are shown in Tables 3 and 4.

Table 3
Collision Data Samples

Physics Sample	Official CMS Datasets
Run 2010A Electron	*/EG/Run2010A-Nov4ReReco_v1/RECO*
Run 2010B Electron	*/Electron/Run2010B-Nov4ReReco_v1/RECO*

Table 4
MC Samples

Physics Sample	Official CMS Datasets
$Z \rightarrow \tau\tau$	*/DYToTauTau_M-20_TuneZ2_7TeV-pythia6 -tauola/Fall10-START38_V12-v1/*
$Z \rightarrow ee$	*/DYToEE_M-20_TuneZ2_7TeV-pythia6/Fal l10-START38_V12-v1/*
$W + jets$	*/WJetsToLNu_TuneZ2_7TeV-madgraph-tauo la/Fall10-START38_V12-v1/*
$t\bar{t}$	*/TT_TuneZ2_7TeV-pythia6-tauola/Fall10 -START38_V12-v1/*
$\gamma + Jets$ $0 \rightarrow 15$	*/G_Pt_0to15_TuneZ2_7TeV_pythia6/Fa ll10-START38_V12-v1/*
$\gamma + Jets$ $15 \rightarrow 30$	*/G_Pt_15to30_TuneZ2_7TeV_pythia6/F all10-START38_V12-v1/*
$\gamma + Jets$ $30 \rightarrow 50$	*/G_Pt_30to50_TuneZ2_7TeV_pythia6/F all10-START38_V12-v1/*
$\gamma + Jets$ $50 \rightarrow 80$	*/G_Pt_50to80_TuneZ2_7TeV_pythia6/F all10-START38_V12-v1/*
$\gamma + Jets$ $80 \rightarrow 120$	*/G_Pt_80to120_TuneZ2_7TeV_pythia6/ Fall10-START38_V12-v1/*
$\gamma + Jets$ $120 \rightarrow 170$	*/G_Pt_120to170_TuneZ2_7TeV_pythia6 /Fall10-START38_V12-v1/*
$\gamma + Jets$ $170 \rightarrow 300$	*/G_Pt_170to300_TuneZ2_7TeV_pythia6 /Fall10-START38_V12-v1/*
$\gamma + Jets$ $300 \rightarrow 470$	*/G_Pt_300to470_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/*
$\gamma + Jets$ $470 \rightarrow 800$	*/G_Pt_470to800_TuneZ2_7TeV_pythia6/Fall10-START38_V12-v1/*
qcd EM $20 \rightarrow 30$	*/QCD_Pt-20to30_EMEnriched_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/*
qcd EM $30 \rightarrow 80$	*/QCD_Pt-30to80_EMEnriched_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/*
qcd EM $80 \rightarrow 170$	*/QCD_Pt-80to170_EMEnriched_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/*

VII.3 Hadronic Tau Reconstruction and Identification

Hadronic taus are characterized by narrow signatures in the detector. This happens as consequence of the kinematics of tau decays. For instance, the momentum of the decay products of a hadronic tau decaying to a charged pion and a neutrino, as shown in Figure 33, is given by (momentum with respect the tau rest frame):

$$\begin{aligned} p_{\pi_\tau} &= \tfrac{m_\tau}{2}(0, sin\beta, cos\beta, 1) \\ p_{\nu_\tau} &= \tfrac{m_\tau}{2}(0, -sin\beta, -cos\beta, 1) \end{aligned} \tag{40}$$

The quantity of interest for us is the angle of the decay products of the boosted tau with respect to the laboratory frame. Therefore, to measure this we need to transform the angle in the tau rest frame (β) to the laboratory frame:

$$\alpha = arccos\Big(\frac{p_\tau^2 sin^2\beta - m_\tau^2}{p_\tau^2 sin^2\beta + m_\tau^2}\Big) \tag{41}$$

The equation above shows that the angle between the decay products of the boosted tau in the laboratory frame depends on the tau transverse momentum. For cases where the tau transverse momentum is high compared to the invariant mass of the tau ($p_{tau} >> m_{\tau}$), the α angle goes to zero and the tau decay products are confined in a narrow cone about the original tau direction.

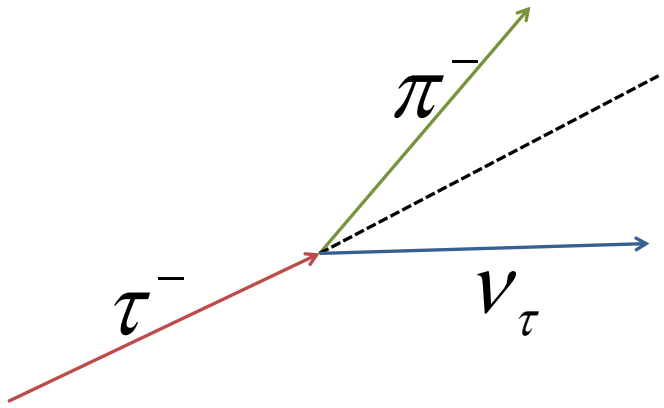
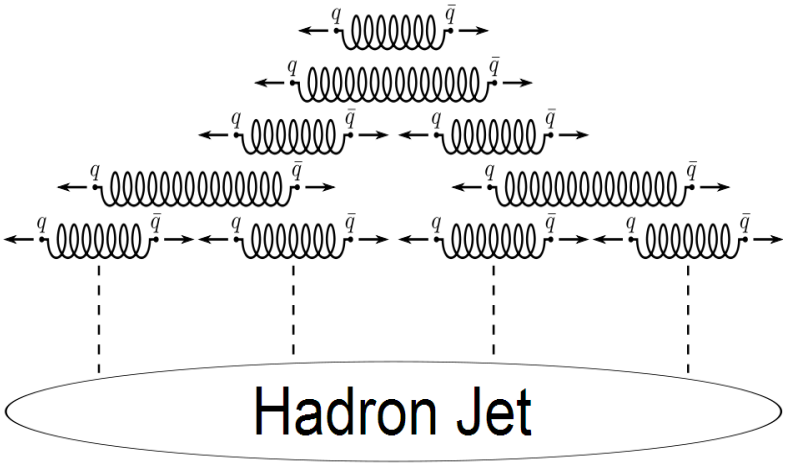


Fig. 33. Sketch of a hadronic tau decay in the boosted tau frame.

Events with direct production of quarks and gluons, known as QCD processes, represent the most significant source of tau jet misidentification. The probability for a jet to look like a narrow tau signature is referred to as the jet to tau fake rate. The production of these types of events has a larger cross-section by several orders of magnitude with respect to events leading to final states with taus. For all practical purposes quarks and gluons do not exist as free particles. Instead they undergo a process known as fragmentation or hadronization where additional quark pairs and gluons are produced [24]. As these quark pairs are pulled apart, the energy of the color field that governs the strong force can be large enough to continue producing quarks and gluons that eventually form bounded states of hadrons. A sketch of this process is shown in Figure 34. Due to the hadronization process jet-like structures



(a)

Fig. 34. Sketch of the Jet hadronization process in QCD events.

characterized with a high multiplicity of hadrons with low energies can be created. Unlike taus, QCD jets have a high multiplicity of particles with soft momentum and tend to have a much broader signature in the detector. The differences are used to discriminate QCD jets from real tau jets.

We use a shrinking cone algorithm for the tau identification, complemented by an elliptical isolation for photons. The shrinking cone is a cone in $\eta - \phi$ space with an annulus of radii $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$, weighted by the transverse energy of the tau jet candidate:

$$\Delta R = \Delta R_0 / E_T \tag{42}$$

where ΔR_0 has a constant value of 5.0, which yields the best efficiency for identification of real tau jets with the lowest fake rate when searching for taus in a broad range of transverse momentum.

The charged pion with the highest transverse momentum for the tau jet candidate (leading pion) is enclosed using the shrinking cone, also known as the signal cone.

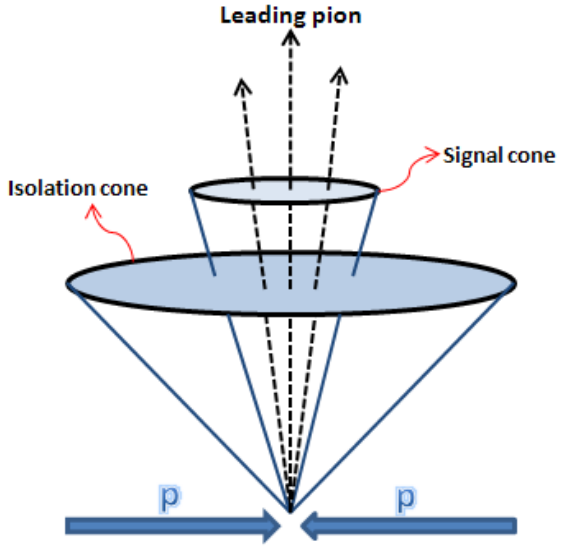
To obtain the highest efficiency for tau identification and the lowest possible fake rate, the signal cone was bounded between $0.07 < \Delta R = 5.0/E_T < 0.15$ [25].

The charged and neutral pions and the other particles in the jets were reconstructed using the Particle Flow (PF) algorithm. This technique reconstructs particles using as much useful information from the CMS detector as possible [25]. For instance, an electron is characterized not only by the momentum reconstructed in the tracking volume but also by the amount and nature of the energy deposited in the calorimetry. We find that the PF algorithm provides more robust identification criteria when compared to other methods that employ a traditional cut approach.

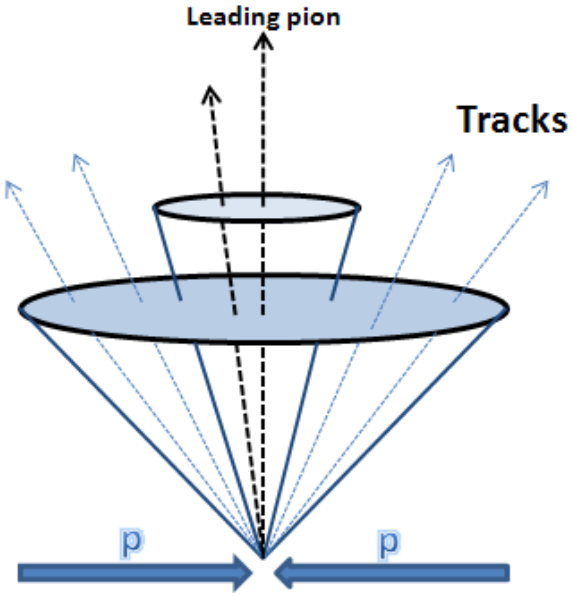
After the leading pion has been identified and enclosed in the signal cone, a larger cone of $\Delta R = 0.5$, known as the isolation cone, is set around signal cone (Figure 35 (a)). The purpose of this larger cone is to discriminate against the QCD fake rate (see Figure 35). As explained above, QCD events are characterized by a high multiplicity of tracks around the jet direction as shown in Figure 35 (b). By setting a threshold for the sum of the transverse momentum of tracks that enter into the isolation region, we suppress dramatically the tau jet misidentification due to QCD events.

In addition to the shrinking cone algorithm, we use an elliptical isolation for photons. Most hadronic tau decays have associated neutral pions (π^0). These π^0 quickly decay to photons which can then interact with the material in the tracking volume and produce electron positron pairs. Because of the strong magnetic field in this volume, these secondary particles can be swept into the isolation region of a traditional cone cut and the tau is lost. In order to recover these taus, we produce a signal cone (see Figure 36) that is longer in the bending plane of the magnet: $R_\phi = 0.15$ and a $R_\eta = 0.07$.

In this analysis isolation is divided in two components, tracks and photons ($\gamma's$). Track isolation is defined as the sum of the reconstructed transverse momentum of all tracks above a set transverse momentum that fall between the signal cone and the isolation cone. Similarly, isolation for photons, known also as Ecal isolation, is defined as the sum of energy depositions in the electromagnetic calorimeter (EcalRecHits) above a set threshold in the transverse energy that fall between the signal and the

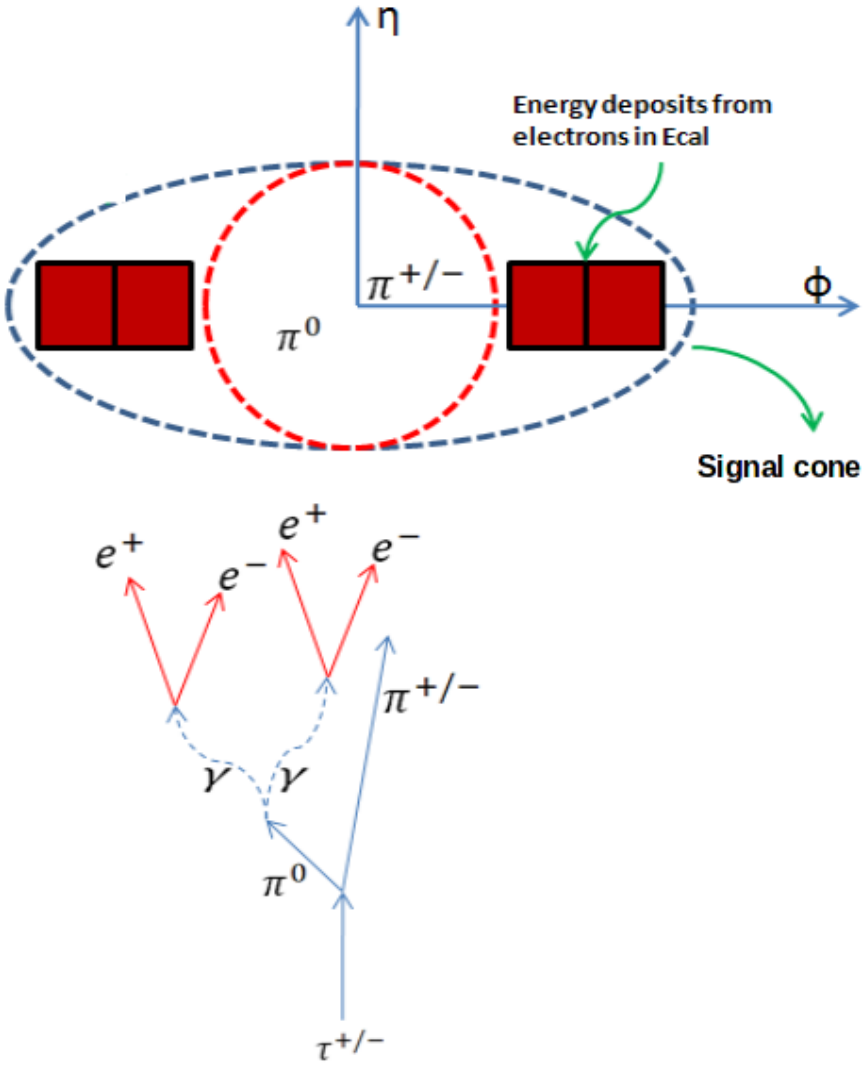


(a)



(b)

Fig. 35. (a) Representation of a three prong hadronic tau decay, identified using the shrinking cone algorithm. (b) Hadronic tau fake rate from QCD events. The narrowness and low multiplicity of tau decays are the main characteristic for their discrimination against QCD events.



(a)

Fig. 36. Representation of the elliptical isolation for EcalRechits. The decay of neutral pions to photons produces electron positron pairs that can be swept out of a circular signal cone region and into the isolation cone. By using an elliptical signal cone, we recover the taus that would otherwise be lost while still removing a majority of the backgrounds.

isolation cones. A mathematical description of isolation is provided in Equations 43 and 44:

$$Track\ Isolation = \sum_{\Delta R_{Iso}=0.5,\ \Delta R_{signal}=\frac{5.0}{E_T}}^{Tracks} p_T < X \tag{43}$$

$$Ecal\ Isolation = \sum_{\Delta R_{Iso}=0.5,\ \Delta R_{signal}^{\gamma' s}}^{\gamma' s} E_T < X \tag{44}$$

In Equation 44 the elliptical signal cone for photons ($\Delta R_{signal}^{\gamma' s}$) is defined as:

$$\Delta R_{signal}^{\gamma' s} = \sqrt{\Big(\frac{\Delta\eta^{\gamma' s}}{R_{\eta}^{Ecal\ signal}}\Big)^2 + \Big(\frac{\Delta\phi^{\gamma' s}}{R_{\phi}^{Ecal\ signal}}\Big)^2} \tag{45}$$

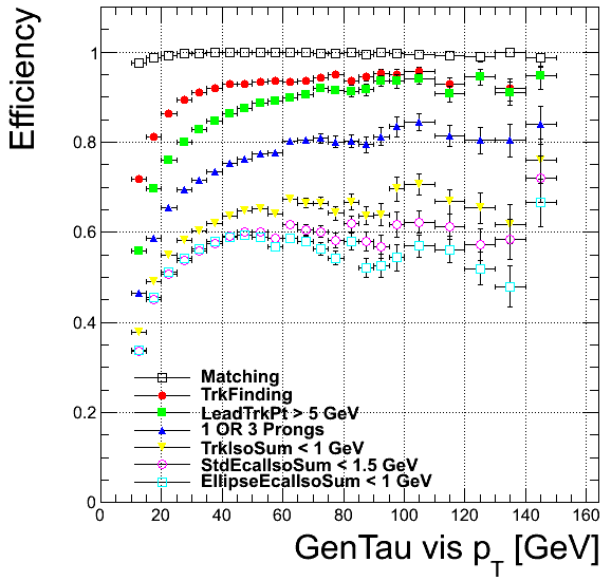
The efficiencies for the tau identification and the fake rate, measured as a function of the transverse momentum of the generated particles for taus and jets, are shown in Figure 37. The efficiency is measured with respect to the true taus that passed each selection criterion with respect to all the tau candidates:

$$\varepsilon^{\tau\ selection} = \frac{Matched\ \tau\ passing\ selection\ x}{All\ \tau\ candidates} \tag{46}$$

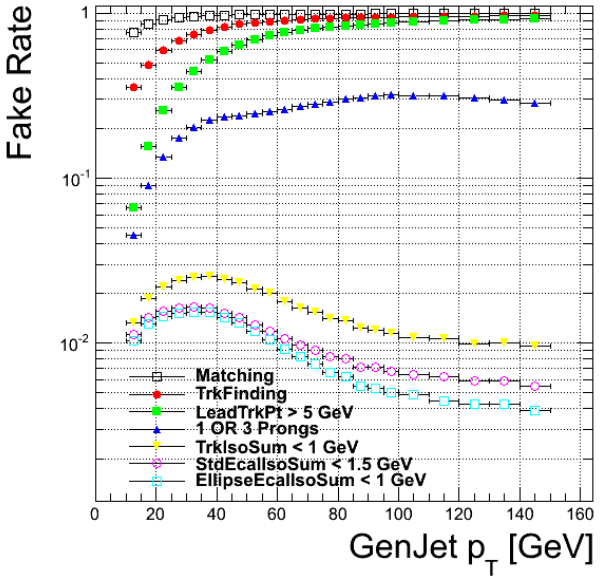
We define a tau to be "matched" as a reconstructed tau in the simulation when it is closely related to a tau at the generator level. The matching is performed by selecting the closest reconstructed tau to a generator level tau (tau generated in the Monte Carlo level), using a cone in $\eta - \phi$ space ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$).

VII.4 Electron Reconstruction and Identification

Electrons in CMS are reconstructed using information from the tracker detector and the electromagnetic calorimeter. Besides leaving hits in the detector material, bremsstrahlung radiation (photons) is produced by electrons passing through the silicon tracker material. Also, electrons produce bremsstrahlung radiation when they bend in the presence of a magnetic field. The energy of the radiated photons



(a)



(b)

Fig. 37. (a) Efficiency for the different steps in the tau identification algorithm as a function of the generated tau visible transverse momentum (p_T) [4]. (b) Fake rate measured as the probability that a generated quark/gluon jet is reconstructed as a PFTau using the shrinking cone algorithm. Notice the dramatic reduction in the fake rate after isolation criteria are applied. The efficiencies are measured for the case where the hadronic tau decays to one or three charged hadrons plus associated neutral pions and a neutrino.

is spread over several crystals of the electromagnetic calorimeter along the electron trajectory, mostly in the ϕ direction (the magnetic field is in the z direction). If the radiated photons have a high enough energy they can create an electromagnetic shower in the calorimeter. Furthermore, the direct interaction of electrons with the material of the electromagnetic calorimeter also creates an electromagnetic shower. For high energy electrons the energy depositions coming from the radiated photons and the direct interaction of the electrons with the electromagnetic calorimeter are expected to be confined in a small region of the detector. Two algorithms based on energy clustering, "Hybrid" for the barrel and "Island" for the endcaps, are used to measure the energy of electrons and photons that reach the surface of the electromagnetic calorimeter [26].

Electron tracks are reconstructed by matching trajectories in the silicon strip tracker to seed hits in the pixel detector. A pixel seed is composed of two pixel hits compatible with the beam spot. A Gaussian Sum Filter (GSF) is used for the reconstruction of trajectories in the silicon strips. In order to minimize the many possible trajectories due to different combinations of hits, the track that best matches an energy supercluster (cluster of clusters) in the electromagnetic calorimeter is chosen to be the reconstructed track.

The preselection of primary electron candidates requires good geometrical matching and good agreement between the momentum of the track and the energy of the supercluster. Two quantities are used to estimate the geometrical matching, $\Delta\eta_{in} = \eta_{sc} - \eta_{vertex}^{Track}$ and $\Delta\phi_{in} = \phi_{sc} - \phi_{vertex}^{Track}$. The η_{sc} and ϕ_{sc} coordinates correspond to the supercluster position and are measured using an energy weighted algorithm. The η_{vertex}^{Track} and ϕ_{vertex}^{Track} coordinates are the position of the track at the interaction vertex extrapolated, as a perfect helix, to the electromagnetic calorimeter. The good energy-momentum matching is measured by taking the ratio between the corrected energy E_{corr} in the supercluster and the momentum of the track P_{in} measured in the inner layers of the tracker.

Because an electron deposits most of its energy in the electromagnetic calorimeter, the energy deposition in the Hadronic Calorimeter (Hcal) is expected to be small.

This characteristic is implemented by requiring the energy ratio between the hadronic and the electromagnetic seed cluster (E_m) to be small (H/E_m) $< X$. The cuts used for the primary selection of electrons are shown on Table 5. These selections have loose values and are only used for pre-identification of electron candidates. We use tighter selections in order to decrease the electron fake rate. The tighter criteria used for electron identification will be described in Section VII.6.2.

Table 5
Primary Selections for Electron Reconstrction.

Cut	Primary Selection
$\lvert\Delta\eta_{in}\rvert$	< 0.02
$\lvert\Delta\phi_{in}\rvert$	< 0.15
H/E_m	< 0.15
E_{corr}/P_{in}	< 3.0

VII.5 Electron Triggering

The selections for the electron trigger changed during the 2010 data taking period to account for increased luminosities. The highest unprescaled non-isolated single electron triggers were used to select candidate events for the $e\tau$ final state. Table 6 shows the trigger paths chosen for the analysis of the $e\tau$ final state.

Table 6
Trigger Paths for Electrons.

Path Name	Trigger Menu	L1 Condition	Prescale
HLT_Ele15_LW_L1R	$1.6E30$	L1_SingleEG5	1
HLT_Ele15_SW_L1R	$3.5E30$	L1_SingleEG5	1
HLT_Ele17_SW_LooseEleId_L1R	$2E31$	L1_SingleEG8	1
HLT_Ele17_SW_TightEleId_L1R	$6E31$	L1_SingleEG8	1
HLT_Ele22_SW_TighterEleId_L1R_v2/v3	$2E32$	L1_SingleEG8	1

Table 7 shows the definitions that correspond to each trigger path name shown on Table 6. The overall trigger efficiency for electrons and the corresponding data-MC scale factors, measured using standard tag and probe methods [28], are shown in Table 8. The tag and probe technique uses simulated data of physics processes such as $Z \rightarrow ee$ or $Z \rightarrow \mu\mu$ in order to measure the efficiency a particular selection criteria used to identify a specific particle (electrons, muons, taus etc). This is accomplished by applying very strong selection criteria to one of the objects of the pair (tag) and loose selection criteria and the condition to study (e.g isolation) to the other (probe).

Table 7
Electron Trigger Definitions.

Name	H/E_m	$\lvert\Delta\eta_{in}\rvert$ Barrel (Endcaps)	$\lvert\Delta\phi_{in}\rvert$	$\sigma_{i\eta i\eta}$ Barrel (Endcaps)
LooseEleId	0.15	—	—-	0.014 (0.035)
TightEleId	0.15	0.01	0.08	0.012 (0.032)
TighterEleId	0.05	0.008 (0.007)	0.1	0.011 (0.031)

Table 8
Electron Trigger Efficiencies (with $p_T > 15$ GeV)

	$0 < \lvert\eta\rvert < 1.479$	$1.479 < \lvert\eta\rvert < 2.5$	$0 < \lvert\eta\rvert < 2.5$
	Efficiencies		
MC(low-p_T)	0.9492 ± 0.0003	0.9483 ± 0.0005	0.9489 ± 0.0003
Data(low-p_T)	0.9722 ± 0.0020	0.9593 ± 0.0040	0.9686 ± 0.0017
MC(high-p_T)	0.9678 ± 0.0003	0.9704 ± 0.0005	0.9685 ± 0.0002
Data(high-p_T)	0.9757 ± 0.0021	0.9713 ± 0.0038	0.9745 ± 0.0018
	Scale Factors		
Data(low-p_T)	1.0082 ± 0.0022	1.0009 ± 0.0039	1.0062 ± 0.0019
Data(high-p_T)	1.0082 ± 0.0022	1.0009 ± 0.0039	1.0062 ± 0.0019

Table 9
Electron Identification Efficiencies (with $p_T > 15$ GeV)

	$0 < \|\eta\| < 1.479$	$1.479 < \|\eta\| < 2.5$	$0 < \|\eta\| < 2.5$
	Efficiencies		
MC(no Trigger)	0.8635 ± 0.0005	0.7276 ± 0.0010	0.8230 ± 0.0005
Data(no Trigger)	0.8352 ± 0.0040	0.7270 ± 0.0001	0.7995 ± 0.0035
MC(low-P_T)	0.8723 ± 0.0004	0.7339 ± 0.0009	0.8309 ± 0.0004
Data(low-P_T)	0.8477 ± 0.0041	0.7248 ± 0.0086	0.8090 ± 0.0044
MC(high-P_T)	0.8647 ± 0.0005	0.7276 ± 0.0010	0.8240 ± 0.0005
Data(high-P_T)	0.8392 ± 0.0057	0.7203 ± 0.0099	0.7998 ± 0.0002
	Scale Factors		
Data(no Trigger)	0.9672 ± 0.0047	0.9992 ± 0.0014	0.9714 ± 0.0043
Data(low-P_T)	0.9718 ± 0.0047	0.9876 ± 0.0118	0.9736 ± 0.0053
Data(high-P_T)	0.9705 ± 0.0066	0.9900 ± 0.0137	0.9706 ± 0.0006

VII.6 Selection Criteria for Electron + Hadronic Tau Final State

The requirements used to select $e\tau$ pairs are factorized in four categories: acceptance, electron identification, τ identification and topological selections.

VII.6.1 Acceptance Selections

Because the Z' is predicted to be a heavy particle, the decay products are expected to be back to back. At acceptance level we require the electron and hadronic tau pair to be separated by a $\Delta R(e, \tau) > 0.7$ in $\eta - \phi$ space. By applying this cut we remove a significant number of events coming from low p_T backgrounds and uncorrelated processes such as $W + jets$ where the direction of the W and the jet are random. We also avoid counting events where a candidate passes simultaneously the primary selections for electrons and taus, which results in pairs formed by the same object.

After this initial selection, we apply a threshold in the transverse momentum for the hadronic taus and the electron candidates. A threshold cut above 20 GeV for taus and 25 GeV for electrons is applied. The values were chosen to be high enough in order to reject as much background as possible and step away from trigger bias

effects, but low enough to keep a control region with $Z \rightarrow \tau\tau$ events to validate our tau identification methodology.

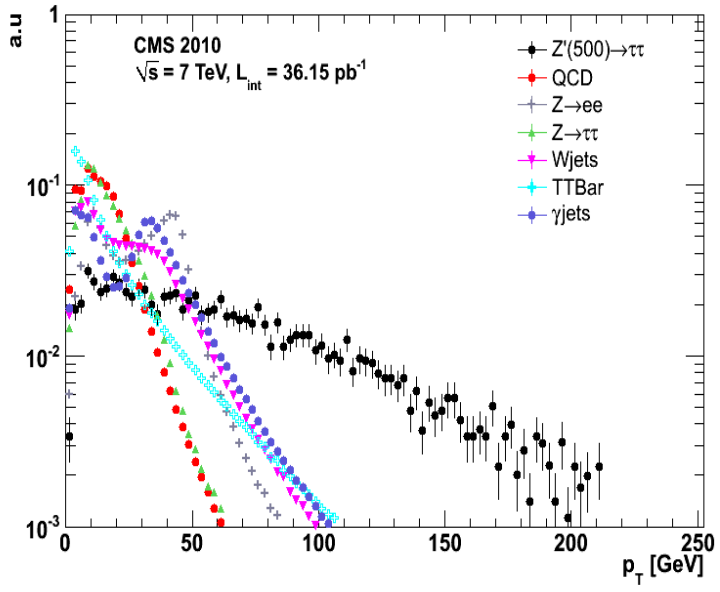
In Section VII.3, we showed that reconstructed tau candidates used good quality tracks in an isolation cone of $\Delta R(\eta, \phi) = 0.5$ around the leading hadron. In order to keep the isolation cone within the tracker acceptance region ($|\eta_\tau| < 2.5$), we constrain the η space for taus to $|\eta_\tau| < 2.1$. To be consistent with the cut applied for taus, we use use the same selection criteria for electrons ($|\eta_e| < 2.1$). Cracks and gaps in the detector acceptance are specifically excluded in the selection. We show some of the kinematic distributions after acceptance selections for signal and background in Figures 38 and 39. There are several features worth noting.

Notice the hard spectrum in the transverse momentum distribution for the Z' sample with respect to the backgrounds. The electron and tau distributions for the $Z \rightarrow ee$ background have a double peak structure; no electron or tau identification criteria has been applied at this stage. The first peak for the electron distribution corresponds to soft jets faking the electron and the second peak corresponds to real electrons coming from the Z. There is no missing energy in the topology of the $Z \rightarrow ee$ decay and the decay electrons, with an expected average momentum of 45 GeV each, make the secondary peak in Figures 38 and 39.

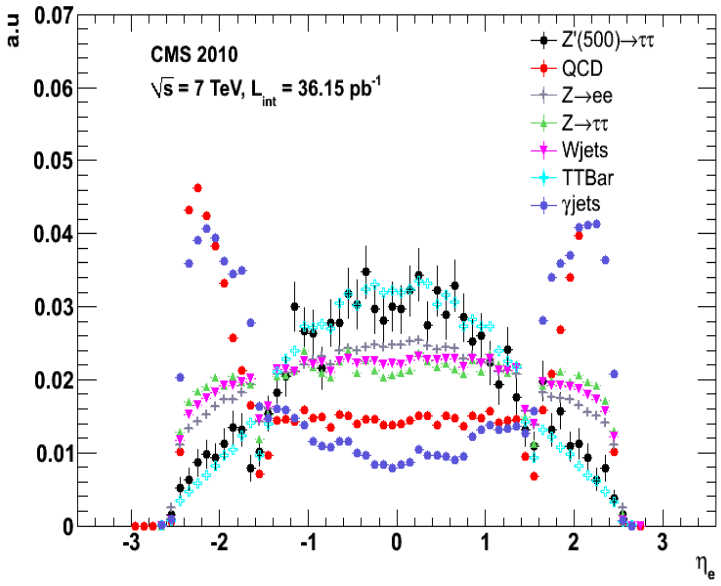
For the $Z \rightarrow \tau\tau$ the transverse momentum spectrum is softer in both legs (electron and tau) with respect to the $Z \rightarrow ee$ decay. This is due to the associated neutrinos in the hadronic (one) and leptonic (two) tau decays. The first peak in the tau momentum distribution corresponds to soft jets faking the hadronic tau. The mean of the electron distribution is around 15 GeV as expected, due to the undetected neutrinos (1/3 of the $M_Z/2$).

The electron transverse momentum distribution for the $W+jets$ sample presents a secondary structure around 40 GeV. This corresponds with the real electrons coming from the W decay. The low part of the spectrum is dominated by low momentum jets faking the electron signature.

The $t\bar{t}$ distribution for the electron momentum peaks at the low part of the spectrum, followed by a long tail. This sample has a high number of jets coming

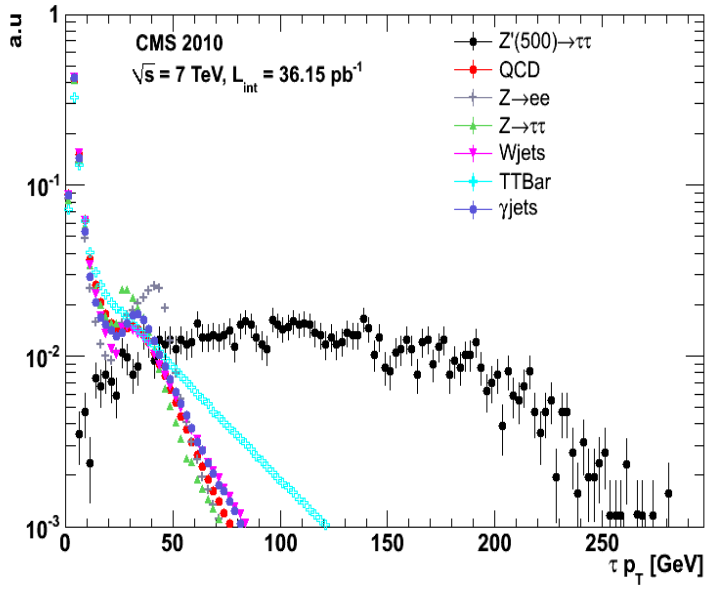


(a)

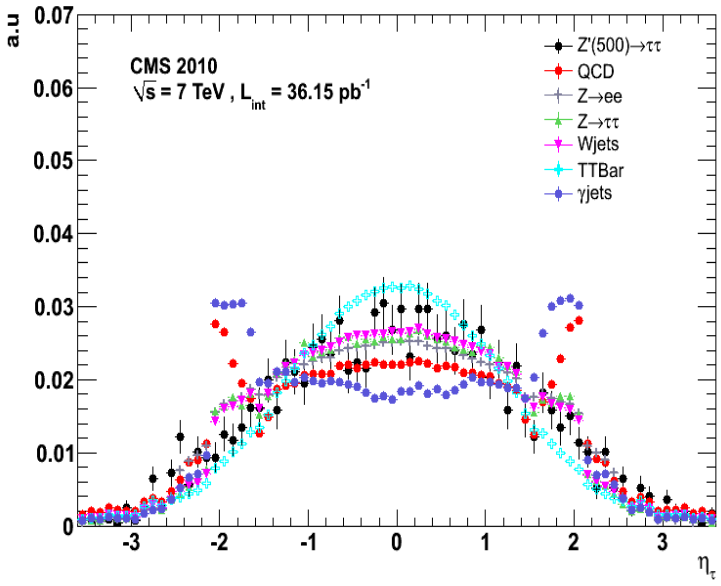


(b)

Fig. 38. Kinematic distributions for electrons. Each successive plot contains all the cuts from the preceding selection. (a) Electron p_T, (b) Electron η.



(a)



(b)

Fig. 39. Kinematic distributions for taus. Each successive plot contains all the cuts from the preceding kinematic selection for electrons. (a) Tau p_T, (d) Tau η.

from the b quark which fake the real electrons from the W decay since no electron identification criteria has been applied at this stage. In Figure 39 (a) the distributions for $W + jets$ and $t\bar{t}$ are almost identical because the hadronic tau is faked by jets or W decays to hadrons or leptons.

Figure 38 (b) and Figure 39 (b) show the η distributions for electrons and taus respectively. Events such as $Z' \rightarrow \tau\tau$, $t\bar{t}$ and $Z \rightarrow ee$ have higher statistics in the central region of the detector ($|\eta| < 1.2$). These type of particles are produced in the hard scattering interaction among quarks. The efficiency drop in $1.4 < |\eta| < 1.6$ is due to cracks in the detector volume between the barrel and the endcaps. Because of the high activity in the forward direction (beam direction), it is more difficult to discriminate and reconstruct good physics events, which results in low efficiency in this sector of the detector acceptance.

Events coming from QCD and $\gamma + Jets$ processes are expected to have a higher number of events in the endcap region, as most of the jets come from soft scattering processes in the beam direction.

After the acceptance selections have been applied, the major background contributions come from QCD events, Drell-Yan processes and events where electrons are created through the decay of a W boson. As described in Sections VII.3 and VII.4, we used selection criteria approved by the CMS Physics Object Group (POG) for electron and tau identification in order to reduce the fake rate of jets and leptons. Nevertheless, we studied carefully each selection in order to corroborate that the parameters and cut values were optimal for our analysis. Furthermore, we measured the efficiency of each cut for the simulated signal samples and all the backgrounds as shown in Tables 11 to 15.

VII.6.2 Electron Identification

Our electron selections have two main components, electron identification (eID) and electron isolation. The electron identification is based on High Energy Electron Pairs (HEEP) approved by the CMS physics object group. The HEEP selections are

a set of selections optimized to identify high energy electrons which are used in several CMS analyses. Before choosing this set of selections for the electron identification, we had our own selection criteria. Nevertheless we decided to use the recommended HEEP selections as they gave us a similar identification power with the advantage of being already approved by the CMS collaboration. Although the HEEP selections were approved by the experiment, we studied each selection separately in order to see if we obtained a good efficiency for signal and low fake rates for the different backgrounds. The HEEP selections are described below.

We use electron candidates identified using energy clusters in the electromagnetic calorimeter. These electrons are known as Ecal driven electrons. Table 10 shows the list of selection criteria applied for the electron identification. In Table 10 $\sigma_{i\eta i\eta}$ is an energy weighted variable useful to estimate the cluster shape. This variable is defined as the sum of the difference in η between the seed crystal in the electromagnetic calorimeter and a neighbor crystal i, which is weighted by their energy ratio (Equation 47).

$$\sigma_{i\eta i\eta} = \sum_{crystals} (\eta_{seed} - \eta_i)^2 * \frac{E_i}{E_{seed}} \tag{47}$$

The variable $E^{n\times m}/E^{m\times m}$ measures the energy ratio between an array of crystals of size $n \times m$ in the electromagnetic calorimenter, where $n < m$, and one of size $m \times m$. These two variables are used to measure the spread of the electromagnetic shower. High energy electrons are characterized by narrow energy depositions in the electromagnetic calorimeter. For this reason we expect that the energy deposited by a high energy electron in a $E^{1\times 5}$ array of crystals around the crystal with the highest energy deposition, should be about the same when compared with a larger array of $E^{5\times 5}$ around the same crystal with the highest energy deposition:

$$\frac{E^{n\times m}}{E^{m\times m}} \approx 1 \tag{48}$$

The definitions of $\Delta\eta_{in}$, $\Delta\phi_{in}$ and H/E_m were explained in Section VII.4. In order to obtain the best possible identification of electrons, the selections applied

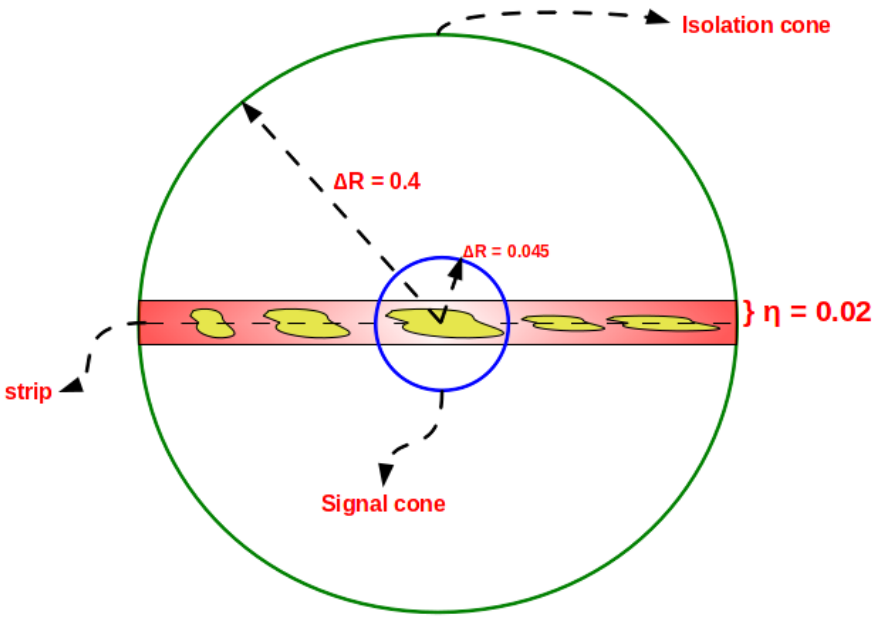
for these variables are tighter than the primary selections used at the reconstruction level.

The interaction of high energy photons with the detector material can produce electron-positron pairs that can fake electrons coming from decays such as $Z \rightarrow ee$. These conversion electrons from photons can also fake the electron coming from the leptonic tau. We remove electrons coming from conversions by requiring that the electron tracks must not have missing hits in the inner layers of the pixel detector, as most conversions happen in the outer layers of the tracker detector material.

Isolation is used for tracks and energy depositions in the electromagnetic calorimeter (EcalRecHits). For the electromagnetic isolation, the energy depositions in the electromagnetic calorimeter across the ϕ direction of the electron candidate are enclosed in a rectangular strip of specified width in η. Additionally, a small cone is placed around the supercluster with the highest energy deposition as shown in Figure 40. The strip and the small signal cone form what we call the electron signal veto. A larger cone is placed around the signal veto region, this cone is called isolation cone. The sizes of the signal strip and cone, and the isolation cone are specified below:

- η_{min} = -0.02, η_{max} = 0.02
- ΔR_{signal} = 0.045
- $\Delta R_{isolation}$ = 0.4
- Threshold for the transverse energy of photons that fall between the signal and the isolation region in the barrel of the detector: 0.08 GeV
- Threshold for the transverse energy of photons that fall between the signal and the isolation region in the endcap of the detector: 0.1 GeV

Under this parameters we defined electromagnetic isolation as the sum of the transverse energy of photons above a threshold, that fall between the signal and the isolation region.



(a)

Fig. 40. Sketch of the strip and signal cone used to enclose the energy depositions of the electron candiate in the electromagnetic calorimeter.

$$eEcal_{iso} = \sum_{\Delta R<0.4} E_T^{Ecal\ RecHits} < X \tag{49}$$

Isolation for tracks uses a small cone, known as signal cone, around the electron track candidate and a larger cone around the signal cone known as the isolation cone:

- $\Delta R_{signal} = 0.015$
- $\Delta R_{isolation} = 0.4$
- Threshold for the transverse momentum of tracks that fall between the signal and the isolation region : 0.7 GeV

The track isolation is defined as the sum of the transverse momentum of tracks above a threshold, that fall between the signal and the isolation region:

$$eTrk_{iso} = \sum_{\Delta R<0.4} p_T^{Trks} \; < \; X \tag{50}$$

MC based distributions for electron identification are shown in Figures 41, 42 and 43. Notice the significant reduction we achieve in the electron fake rate, especially for backgrounds dominated by jets such as QCD and $t\bar{t}$, by applying the different electron identification selections. Selections such as $|\Delta\eta_{in}|$, H/E_m and isolation are especially powerful in removing most of the electron fake rate.

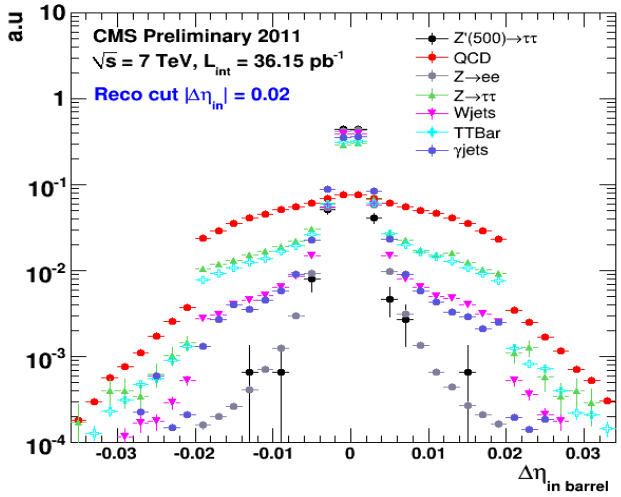
Table 10
Electron ID Selections.

Cut	HEEP (barrel)	HEEP (endcap)
E_T	> 25 GeV	> 25 GeV
isEcalDriven	1	1
$\|\Delta\eta_{in}\|$	< 0.005	< 0.007
$\|\Delta\phi_{in}\|$	< 0.09	< 0.09
H/E_m	< 0.05	< 0.05
$\sigma_{i\eta i\eta}$	n/a	< 0.03
$E^{n\times m}/E^{m\times m}$	$E^{2\times 5}/E^{5\times 5} > 0.94$ || $E^{1\times 5}/E^{5\times 5} > 0.83$	
e Ecal Iso	< 4.5	< 4.5 GeV
e Track Iso	< 3.5	< 3.5 GeV

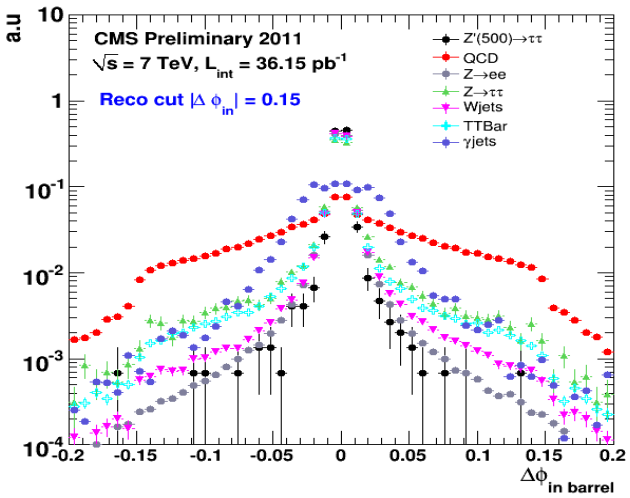
VII.6.3 Tau Identification

For hadronic taus, we require one charged hadron which corresponds to 84% of the hadronic tau decays. Even though we lose 16% of the hadronic tau decays by not using the tau decays with 3 or 5 charged hadrons, we suppress dramatically the fake rate of events coming from QCD processes.

Since tau decays are expected to be well collimated, we require the electron and tau jet candidates to be fairly well isolated as described in Section VII.3. Figure 44 shows the electromagnetic and track isolation distributions for taus.

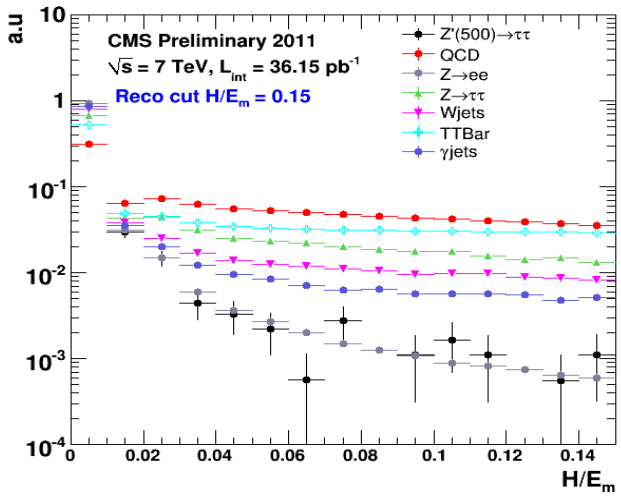


(a)

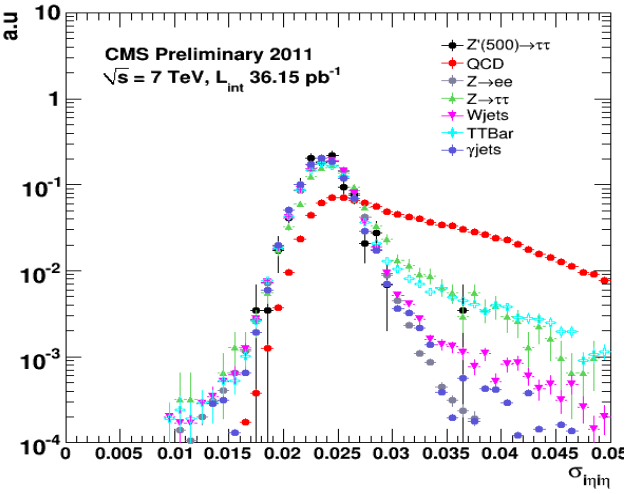


(b)

Fig. 41. Distributions used for the HEEP electron identification. Each successive plot contains all the cuts from the preceding plots. The distributions are normalized to one. (a) $\Delta\eta_{in\ barrel}$, (b) $\Delta\phi_{in\ barrel}$. A preselection cut at $|\Delta\eta_{in\ barrel}| = 0.02$ and $|\Delta\phi_{in\ barrel}| = 0.15$ is applied at the reconstruction level as explained in Section VII.4. Notice much broader distribution for QCD events unlike events with real electrons such as $Z \rightarrow \tau\tau$, $W + Jets$ and our expected Z' signal.

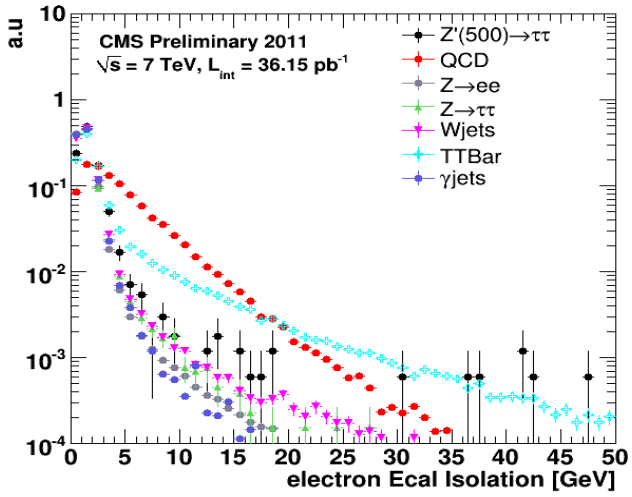


(a)

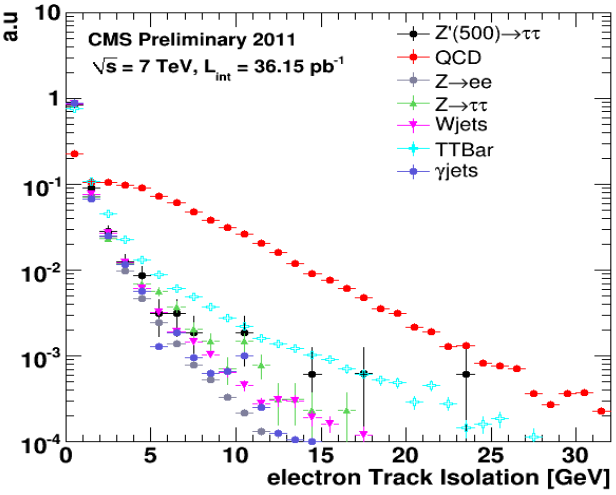


(b)

Fig. 42. Distributions used for the HEEP electron identification. Each successive plot contains all the cuts from the preceding plots. The distributions are normalized to one. (a) H/E_m, the peak on the left of the H/E_m distribution are mostly the real electrons in the event from the different processes. The tail of the distribution corresponds to jets faking the electron. Notice the nice reduction in the electron fake rate that we achieve by cutting at low values of the H/E_m distribution. (b) $\sigma_{i\eta i\eta}$, this variable help us to remove contamination from QCD events. Notice that by cutting below 0.03 we significantly reduce events from QCD.

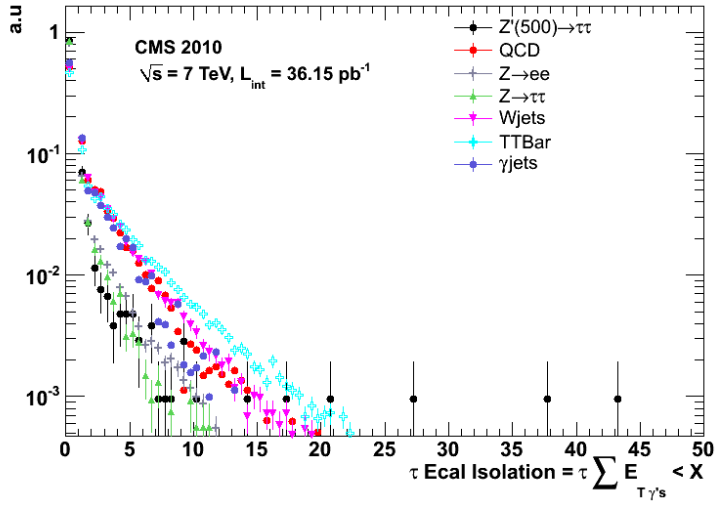


(a)

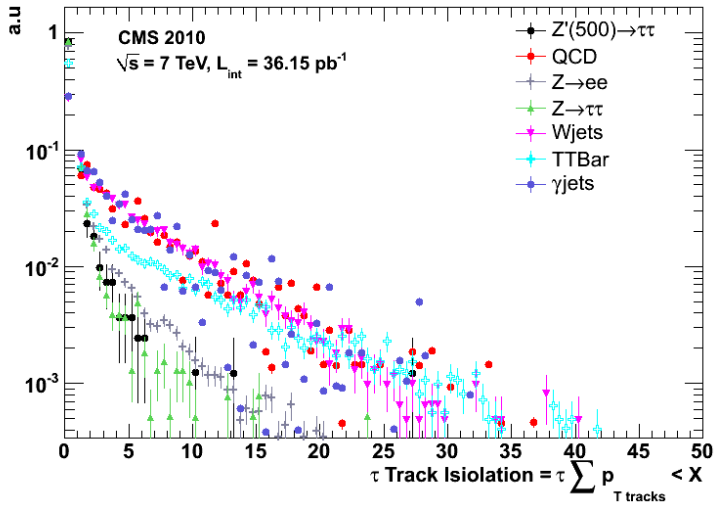


(b)

Fig. 43. Isolation distributions used for the electron identification. (a) Electromagnetic isolation, (b) track isolation. Backgrounds dominated by the presence of jets such as QCD and $t\bar{t}$ can be highly reduced by applying isolation.



(a)



(b)

Fig. 44. Isolation distributions for the hadronic τ, (a) Ecal isolation, (b) track isolation. The distributions are after acceptance, electron and tau ID selection criteria have been applied. Notice that by cutting below 1 GeV we keep most of our signal and suppress a large portion of events with jets such as $W + Jets$, $t\bar{t}$ and QCD.

In order to remove badly reconstructed electrons that fake the hadronic tau jet we implemented an electron veto cut. The electron veto is a combination of two cuts. The first cut uses the ratio between the energy deposited by the tau jet candidate in a 3×3 array of crystals in the hadronic calorimeter and the momentum of the leading charged hadron ($H_{3\times3}/p_{Lp}$). For hadronic taus we expect the charged pions to deposit most of their energy in the hadronic calorimeter and to have well reconstructed tracks with a high quality estimate of the momentum. Because electrons interact only electromagnetically their energy depositions in the hadronic calorimeter are expected to be small. For this reason, by requiring the $H_{3\times3}/p_{Lp}$ ratio to be above a value X we can discriminate a good portion of electrons faking hadronic taus. Additionally, these electrons often have bad quality tracks associated with them. By applying a cut in the number of hits of the leading track for the tau candidate ($N_{hits} > 12$) we were able to suppress badly reconstructed electrons faking the hadronic tau sitting in the high mass region. By applying these two cuts we managed to remove around 96% of the badly reconstructed electrons faking the hadronic tau in the high mass region where we expect to observe our signal.

Figure 45 shows the distribution of the number of hits for the hadronic tau leg in the Z' sample, for the showering electrons, which tend to leave a swath of energies in the electromagnetic calorimeter, from $Z \rightarrow ee$ decays and for high quality, golden, electrons from $Z \rightarrow ee$ with good matching between the track candidate and the seed cluster in the electromagnetic calorimeter. Notice the good separation between the showering electrons, which can become fake taus, and the hadronic taus expected from Z'. The golden electrons provide a good cross check that uses the data rather than the simulation.

VII.6.4 Topological Selections

Since the Z' is expected to be produced almost at rest in the laboratory frame, we require the decay products of the tau pair to be back to back. Due to the

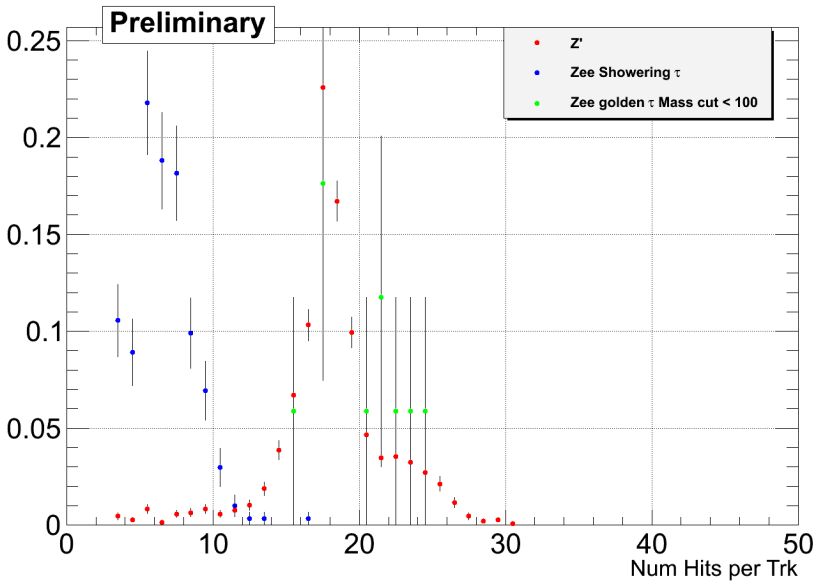


Fig. 45. Distributions of the number of hits for the leading track of the hadronic tau. Notice that the badly reconstructed electrons faking the hadronic tau (showering electrons) have a low number of hits in the tracker. For this reason the momentum of these particles is badly measured which usually leads to an overestimation of the real value. The distributions are normalized to unity.

undetected neutrinos in the tau decays we require missing transverse energy ($\not{E}_{\mathrm{T}}$) above 30 GeV. The threshold was chosen high enough to suppress the backgrounds as much as possible, but low enough to still keep events from $Z \rightarrow \tau\tau$ in order to validate our tau identification criteria. After applying this selection, the most significant backgrounds are $W + Jets$ and $t\bar{t}$. In order to suppress $W + Jets$ events we use a cut known as p_ζ which takes the projection of the missing transverse energy onto a bisector vector traced between the visible products of the two candidates of the pair. Figure 46 illustrates the p_ζ cut.

This cut is very efficient for signal due to the narrow profile of taus with a high transverse momentum. Equation 51 shows how the p_ζ variable is calculated using the projections of the visible momentum and the missing transverse energy onto the bisector axis.

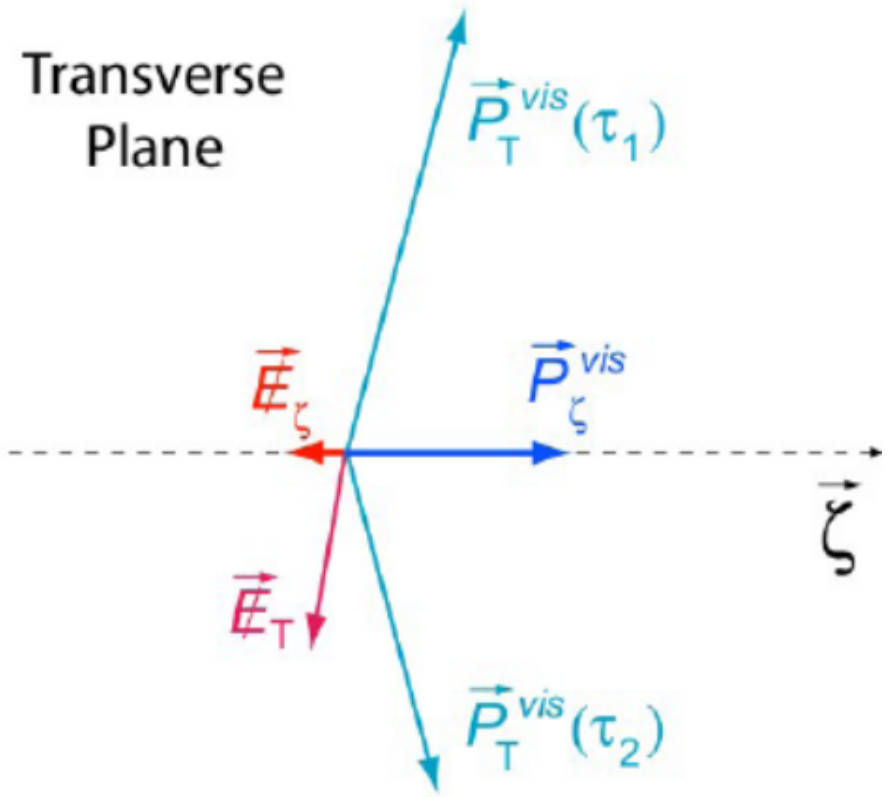
$$\begin{aligned} p_{\zeta\ t}^{vis} &= \overrightarrow{p}^{vis}_{\tau_1\ t} \cdot \hat{\zeta} + \overrightarrow{p}^{vis}_{\tau_2\ t} \cdot \hat{\zeta} \\ p_{\zeta\ t} &= p_{\zeta\ t}^{vis} + \overrightarrow{\not{E}_{\mathrm{T}}} \cdot \hat{\zeta} \end{aligned} \tag{51}$$

Figure 47 shows the p_{zeta} versus p_{zeta}^{vis} for Z' and $W + Jets$ MC events. We apply a two dimensional cut:

$$p_\zeta - 0.875 \times p_\zeta^{vis} > -7 \tag{52}$$

where the cutting values have been chosen in order to obtain a high efficiency for Z' events and reduce events from $W + jets$ as much as possible.

In order to further suppress $t\bar{t}$ events, we use an algorithm to tag jets coming from b quarks (b-Tagging). The algorithm was developed by the CMS b-Tagging group and was officially approved by the physics object group [30]. We use the Track Counting High Efficiency algorithm with Medium working point (TCHEM). The algorithm determines the number of tracks, N, with a significance of the impact parameter ($d_0^{sig} = \frac{d_0}{\sigma(d_0)}$) greater than a set value called the working point. Jets coming from b quarks are expected to be very energetic resulting in a large fragmentation



(a)

Fig. 46. Illustration of the visible momentum and missing transfer energy ($\not\!\!E_{\rm T}$) projection onto a bisector vector traced between the visible products of a two body particle decay: $\overrightarrow{p}_T^{vis}(\tau_1)$ and $\overrightarrow{p}_T^{vis}(\tau_2)$ are the visible transverse momentum of the candidates in the tau pair, $\not\!\!E_{\rm T}$ is the missing transverse energy due to the undetected neutrinos and $\overrightarrow{p}_\zeta^{vis}$ and $\not\!\!E_{{\rm T}\zeta}$ are the projections of the visible transverse momentum and missing transverse energy of the taus onto the bisector vector. Events coming from $W + Jets$ processes are significantly suppressed by requiring a small value of missing transverse energy on the bisector axis, whereas high momentum taus are unaffected due to the narrow profile of their decay products [29].

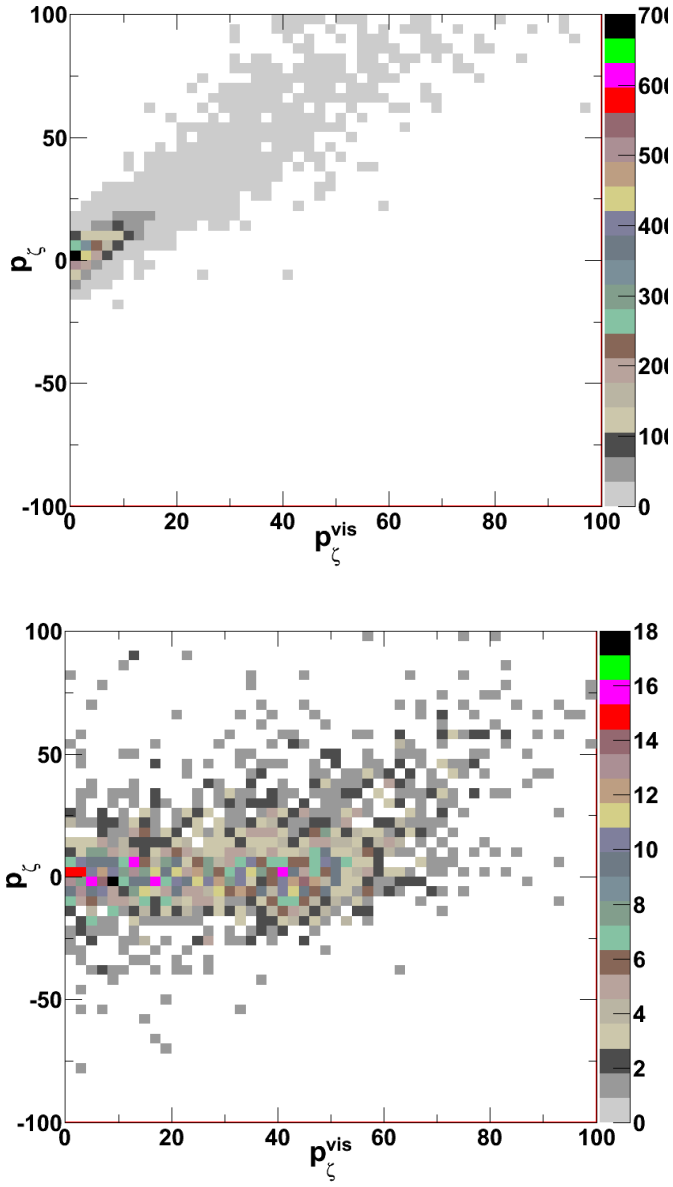


Fig. 47. Distribution for p_ζ vs. p_ζ^{vis} for (a) $Z' \to \tau\tau$, and (b) $W + Jets$. We apply a two dimensional cut in our final selection criteria: $p_\zeta - 0.875 \times p_\zeta^{vis} > -7$. The cutting values were chosen in order to keep a high efficiency for Z' and reduce $W + jets$ events as much as possible.

that produces many subsequent tracks. Besides their large multiplicity, b jets are also characterized by a long lifetime component. For these reasons, a simple track counting algorithm presents a good mechanism for identification of b jets with respect to narrow tau jets with low multiplicity of tracks and small impact parameter. We required zero jets in our selections tagged as $b - jets$.

VII.6.5 Final List of Selection Criteria

All the selection criteria are summarized below:

Acceptance Selection:

- ≥ 1 e with $|\eta| < 2.1, E_T > 25$ GeV
- ≥ 1 PFTau with $|\eta| < 2.1, p_T > 20$ GeV, and leading track with $p_T > 5$ GeV
- $\Delta R(e, \tau) > 0.7$ GeV

e Identification:

- Electron must be a Ecal driven electron
- $|\Delta\eta_{in}| < 0.005$ in EB; $|\Delta\eta_{in}| < 0.007$ in EE
- $|\Delta\phi_{in}| < 0.09$ in EB and EE
- $H/E_m < 0.05$ in EB and EE
- $\sigma_{i\eta i\eta} < 0.03$ in EE
- $E^{2\times5}/E^{5\times5} > 0.94$ $||E^{1\times5}/E^{5\times5} > 0.83$ in EB
- Ecal Isolation : $\sum E_T^{ecal} < 4.5$ GeV
  ($E_{barrel}^{EcalRecHit} > 0.08$ GeV, $E_{T\ endcap}^{EcalRecHit} > 0.1$ GeV, $\Delta R_{iso} = 0.4$)
- Track Isolation: $\sum p_T^{trk} < 3.5$ GeV ($p_T^{trk} > 0.7$ GeV, $\Delta R_{iso} = 0.4$)

τ_h Identification:

- $H_{3x3}/P > 0.03$ and τ seed track number of hits > 11
- Exactly 1 signal charged hadron ($p_T^{chad} > 1$ GeV, $\Delta R_{sig} = 5/E_T$)
- Ecal Isolation : $\sum E_T^{gammas} < 1.0$ GeV ($E_T^{gammas} > 1.0$ GeV, $R_\eta = 0.07$, $R_\phi = 0.15$)
- Track Isolation : $\sum p_T^{chad} < 1.0$ GeV ($p_T^{chad} > 1$ GeV, $\Delta R_{iso} = 0.5$, $\Delta R_{TrackSignal} = 5/E_T$)

Topological requirements:

- $\cos\Delta\phi(e,\tau) <$ -0.95 (τ jet direction is calculated using the sum of the 4-momentum of signal cone constituents)
- $Q(e) \times Q(\tau) < 0$ (τ charge is defined as the charge of the leading track)
- $\not{E}_{\mathrm{T}} > 30$
- $p_\zeta - 0.875 \times p_\zeta^{vis} > -7$
- 0 jets tagged as b-jets

The relative efficiencies for the selection criteria in the $e\tau$ channel are shown in Tables [11 - 15].

Table 11
$e\tau$ $Z^{'}$ Relative Cut Efficiencies (%)

Cut/Selection	Z'_{350}	Z'_{400}	Z'_{500}	Z'_{600}
$p_t^e > 25$	69.62 ± 0.79	72.65 ± 0.77	77.04 ± 0.71	79.90 ± 0.66
$\|\eta^e\| < 2.1$	91.84 ± 0.57	92.07 ± 0.55	93.96 ± 0.46	95.14 ± 0.39
$p_t^\tau > 20$	94.67 ± 0.49	95.65 ± 0.43	96.61 ± 0.36	98.16 ± 0.25
$\|\eta^\tau\| < 2.1$	74.56 ± 0.97	78.24 ± 0.89	78.69 ± 0.83	80.93 ± 0.75
$\tau_{seed} p_T > 5GeV$	91.65 ± 0.71	91.70 ± 0.67	93.70 ± 0.56	94.20 ± 0.49
e IsEcalDriven	98.70 ± 0.30	98.71 ± 0.29	99.22 ± 0.21	99.01 ± 0.22
e $H/E < 0.05$	99.12 ± 0.25	98.82 ± 0.28	98.93 ± 0.24	98.80 ± 0.24
e $\Delta\eta_{\text{in}}$	98.08 ± 0.37	99.01 ± 0.25	99.14 ± 0.22	98.93 ± 0.23
e $\Delta\phi_{\text{in}}$	99.02 ± 0.27	99.13 ± 0.24	99.65 ± 0.14	99.66 ± 0.13
e $\sigma_{i\eta i\eta}$	100.00 ± 0.00	99.86 ± 0.09	99.94 ± 0.06	100.00 ± 0.00
e E^{nxm}/E^{mxm}	99.01 ± 0.27	99.19 ± 0.23	99.13 ± 0.22	98.82 ± 0.24
e MissHits < 1	96.08 ± 0.54	95.91 ± 0.52	95.92 ± 0.48	96.27 ± 0.42
e EcalIso < 4.5	95.44 ± 0.59	96.16 ± 0.51	95.81 ± 0.49	95.35 ± 0.49
e TrkIso < 3.5	97.74 ± 0.43	97.04 ± 0.46	97.34 ± 0.40	97.51 ± 0.36
τ EVeto	91.60 ± 0.81	89.41 ± 0.85	91.33 ± 0.71	92.39 ± 0.62
τN Prongs $= 1$	71.25 ± 1.38	72.04 ± 1.31	73.16 ± 1.18	75.06 ± 1.06
τ EcalIso < 1	81.60 ± 1.40	76.33 ± 1.46	78.34 ± 1.29	77.50 ± 1.18
τ TrackIso < 1	82.13 ± 1.54	84.50 ± 1.42	84.68 ± 1.27	84.50 ± 1.16
$\cos\Delta\phi(e,\tau) < -0.95$	88.43 ± 1.42	89.54 ± 1.31	90.88 ± 1.10	91.20 ± 0.99
$Q(e) * Q(\tau_{seed}) < 0$	99.33 ± 0.38	98.16 ± 0.60	97.57 ± 0.62	97.72 ± 0.55
$\not{E}_{\text{T}} > 30$	69.42 ± 2.18	69.52 ± 2.10	73.30 ± 1.80	79.01 ± 1.51
p_ζ	88.10 ± 1.83	91.89 ± 1.49	89.37 ± 1.46	90.10 ± 1.24
b-Tagging	99.60 ± 0.40	98.52 ± 0.73	99.72 ± 0.28	99.14 ± 0.43

Table 12
$e\tau$ Relative Cut Efficiencies (%)

Cut/Selection	Z$\rightarrow \tau\tau$	Z$\rightarrow ee$	$t\bar{t}$	$W + Jets$
$p_t^e > 25$	26.05 ± 0.09	68.95 ± 0.05	51.08 ± 0.06	58.41 ± 0.05
$\|\eta^e\| < 2.1$	86.51 ± 0.14	93.87 ± 0.03	96.27 ± 0.03	87.69 ± 0.04
$p_t^\tau > 20$	68.63 ± 0.20	93.52 ± 0.03	99.10 ± 0.02	53.66 ± 0.07
$\|\eta^\tau\| < 2.1$	82.18 ± 0.20	79.26 ± 0.05	97.33 ± 0.03	82.03 ± 0.07
$\tau_{seed} p_T > 5GeV$	89.37 ± 0.18	94.81 ± 0.03	95.12 ± 0.04	88.13 ± 0.07
e IsEcalDriven	91.12 ± 0.17	99.30 ± 0.01	85.45 ± 0.06	95.39 ± 0.05
e $H/E < 0.05$	83.21 ± 0.24	99.30 ± 0.01	72.57 ± 0.08	91.97 ± 0.06
e $\Delta\eta_{\text{in}}$	75.88 ± 0.30	98.21 ± 0.02	80.57 ± 0.09	92.93 ± 0.06
e $\Delta\phi_{\text{in}}$	94.97 ± 0.17	99.11 ± 0.01	96.56 ± 0.04	98.46 ± 0.03
e $\sigma_{i\eta i\eta}$	98.23 ± 0.11	99.89 ± 0.00	99.00 ± 0.02	99.49 ± 0.02
e E^{nxm}/E^{mxm}	91.78 ± 0.23	99.58 ± 0.01	91.34 ± 0.07	97.71 ± 0.04
e MissHits < 1	95.06 ± 0.19	98.10 ± 0.02	96.27 ± 0.05	96.22 ± 0.05
e EcalIso < 4.5	97.26 ± 0.15	98.79 ± 0.02	85.12 ± 0.09	96.87 ± 0.04
e TrkIso < 3.5	97.18 ± 0.15	99.20 ± 0.01	93.82 ± 0.07	97.86 ± 0.04
τ EVeto	82.09 ± 0.35	17.54 ± 0.06	93.22 ± 0.07	86.87 ± 0.09
τN Prongs $= 1$	52.11 ± 0.51	41.61 ± 0.18	31.77 ± 0.14	16.25 ± 0.11
τ EcalIso < 1	72.45 ± 0.63	74.00 ± 0.25	35.26 ± 0.26	30.15 ± 0.33
τ TrackIso < 1	84.16 ± 0.60	76.42 ± 0.28	55.47 ± 0.46	26.37 ± 0.57
$\cos\Delta\phi(e,\tau) < -0.95$	68.91 ± 0.83	68.31 ± 0.35	15.67 ± 0.45	23.99 ± 1.08
$Q(e) * Q(\tau_{seed}) < 0$	98.35 ± 0.27	97.91 ± 0.13	96.20 ± 0.60	84.99 ± 1.85
$\not{E}_{\text{T}} > 30$	4.119 ± 0.43	0.338 ± 0.05	83.04 ± 1.21	28.71 ± 2.54
p_ζ	86.05 ± 3.7	71.05 ± 7.30	51.00 ± 1.77	50.55 ± 5.24
b-Tagging	100.00 ± 0.00	100.00 ± 0.00	30.06 ± 2.43	100.00 ± 0.00

Table 13
$e\tau$ Relative Cut Efficiencies (QCD and $\gamma + Jets$) (%)

Cut/Selection	$QCD_{Em20\to30}$	$QCD_{Em30\to80}$	$QCD_{Em80\to170}$
$p_t^e > 25$	5.08 ± 0.01	21.55 ± 0.01	49.84 ± 0.11
$\|\eta^e\| < 2.1$	68.88 ± 0.11	70.83 ± 0.03	76.09 ± 0.12
$p_t^\tau > 20$	25.5 ± 0.13	71.76 ± 0.03	95.45 ± 0.09
$\|\eta^\tau\| < 2.1$	58.48 ± 0.28	75.08 ± 0.03	81.22 ± 0.15
$\tau_{seed} p_T > 5GeV$	63.86 ± 0.36	82.94 ± 0.03	90.36 ± 0.15
e IsEcalDriven	83.27 ± 0.35	87.07 ± 0.03	77.99 ± 0.21
e $H/E < 0.05$	79.13 ± 0.42	60.21 ± 0.05	43.79 ± 0.18
e $\Delta\eta_{\text{in}}$	43.44 ± 0.58	38.01 ± 0.07	37.60 ± 0.18
e $\Delta\phi_{\text{in}}$	75.93 ± 0.75	77.28 ± 0.09	80.79 ± 0.11
e $\sigma_{i\eta i\eta}$	83.8 ± 0.75	80.27 ± 0.10	81.33 ± 0.06
e E^{nxm}/E^{mxm}	69.16 ± 1.02	50.35 ± 0.14	40.60 ± 0.10
e MissHits < 1	66.53 ± 1.25	75.18 ± 0.17	79.35 ± 0.38
e EcalIso < 4.5	87.55 ± 1.08	64.4 ± 0.22	40.91 ± 0.19
e TrkIso < 3.0	68.04 ± 1.62	49.14 ± 0.29	41.02 ± 0.17
τ EVeto	82.32 ± 1.61	85.08 ± 0.30	91.35 ± 0.40
τN Prongs $= 1$	12.8 ± 1.56	15.45 ± 0.33	17.77 ± 0.50
τ EcalIso < 1	47.46 ± 6.50	28.74 ± 1.04	18.42 ± 1.35
τ TrackIso< 1	39.29 ± 9.23	26.29 ± 1.89	22.32 ± 2.28
$\cos\Delta\phi(e,\tau) < -0.95$	72.73 ± 13.43	50.35 ± 4.18	28.00 ± 6.35
$Q(e) * Q(\tau_{seed}) < 0$	62.5 ± 17.12	56.94 ± 5.83	57.14 ± 18.28
$\not{E}_{\text{T}} > 30$	0	0	0
p_ζ	0	0	0
b-Tagging	0	0	0

Table 14
$e\tau$ Relative Cut Efficiencies ($\gamma + Jets$) (%)

Cut/Selection	$\gamma + X_{30\to50}$	$\gamma + X_{50\to80}$	$\gamma + X_{80\to120}$	$\gamma + X_{120\to170}$
$p_t^e > 25$	75.65 ± 0.14	77.62 ± 0.11	76.67 ± 0.09	75.28 ± 0.08
$\lvert\eta^e\rvert < 2.1$	71.49 ± 0.17	74.42 ± 0.12	78.20 ± 0.10	82.43 ± 0.09
$p_t^\tau > 20$	65.67 ± 0.22	91.51 ± 0.09	97.80 ± 0.04	99.22 ± 0.02
$\lvert\eta^\tau\rvert < 2.1$	77.45 ± 0.24	77.88 ± 0.15	80.76 ± 0.11	84.91 ± 0.09
$\tau_{seed} p_T > 5GeV$	86.10 ± 0.22	84.27 ± 0.15	82.52 ± 0.12	82.98 ± 0.10
e IsEcalDriven	38.15 ± 0.34	39.43 ± 0.21	43.13 ± 0.17	45.09 ± 0.15
e $H/E < 0.05$	98.47 ± 0.14	92.84 ± 0.18	83.87 ± 0.19	77.60 ± 0.19
e $\Delta\eta_{\text{in}}$	94.00 ± 0.27	93.23 ± 0.18	89.65 ± 0.17	87.65 ± 0.17
e $\Delta\phi_{\text{in}}$	97.22 ± 0.19	97.85 ± 0.11	97.70 ± 0.09	97.88 ± 0.08
e $\sigma_{i\eta i\eta}$	99.43 ± 0.09	99.44 ± 0.06	99.15 ± 0.05	99.05 ± 0.05
e E^{nxm}/E^{mxm}	98.84 ± 0.13	98.24 ± 0.10	97.27 ± 0.10	96.12 ± 0.11
e MissHits < 1	41.25 ± 0.59	42.05 ± 0.38	42.22 ± 0.31	42.71 ± 0.28
e EcalIso < 4.5	98.07 ± 0.25	97.35 ± 0.19	95.86 ± 0.19	93.78 ± 0.21
e TrkIso < 3.0	98.28 ± 0.24	97.93 ± 0.17	97.72 ± 0.14	97.41 ± 0.14
τ EVeto	83.93 ± 0.70	87.52 ± 0.40	90.90 ± 0.29	91.72 ± 0.25
τN Prongs $= 1$	15.40 ± 0.74	18.51 ± 0.50	17.26 ± 0.39	12.63 ± 0.31
τ EcalIso < 1	38.40 ± 2.56	27.58 ± 1.35	22.24 ± 1.04	19.65 ± 1.05
τ TrackIso< 1	30.22 ± 3.90	19.54 ± 2.28	15.01 ± 1.90	20.14 ± 2.38
$\cos\Delta\phi(e,\tau) < -0.95$	73.81 ± 6.78	61.02 ± 6.35	67.92 ± 6.41	66.67 ± 6.24
$Q(e) * Q(\tau_{seed}) < 0$	48.39 ± 8.98	44.44 ± 8.28	52.78 ± 18.32	47.37 ± 18.10
$\not{E}_{\text{T}} > 30$	0	6.25 ± 6.05	5.263 ± 5.12	5.556 ± 5.40
p_ζ	0	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00
b-Tagging	0	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00

Table 15
$e\tau$ Relative Cut Efficiencies ($\gamma + Jets$) (%)

Cut/Selection	$\gamma + X_{170\to300}$	$\gamma + X_{300\to470}$	$\gamma + X_{470\to800}$
$p_t^e > 25$	73.38 ± 0.08	68.85 ± 0.08	64.44 ± 0.03
$\lvert\eta^e\rvert < 2.1$	88.09 ± 0.07	96.00 ± 0.04	98.73 ± 0.03
$p_t^\tau > 20$	99.67 ± 0.01	99.86 ± 0.01	99.93 ± 0.02
$\lvert\eta^\tau\rvert < 2.1$	89.43 ± 0.07	94.25 ± 0.05	95.94 ± 0.04
$\tau_{seed} p_T > 5GeV$	83.64 ± 0.09	83.81 ± 0.08	83.56 ± 0.03
e IsEcalDriven	43.74 ± 0.13	39.95 ± 0.11	37.27 ± 0.05
e $H/E < 0.05$	73.35 ± 0.17	70.29 ± 0.17	67.35 ± 0.06
e $\Delta\eta_{\text{in}}$	87.28 ± 0.15	86.47 ± 0.15	85.16 ± 0.09
e $\Delta\phi_{\text{in}}$	97.99 ± 0.07	98.15 ± 0.06	97.97 ± 0.12
e $\sigma_{i\eta i\eta}$	99.21 ± 0.04	99.24 ± 0.04	99.25 ± 0.13
e E^{nxm}/E^{mxm}	95.80 ± 0.10	95.40 ± 0.10	94.73 ± 0.19
e MissHits < 1	43.16 ± 0.25	44.24 ± 0.24	45.69 ± 0.24
e EcalIso < 4.5	91.17 ± 0.21	74.17 ± 0.32	28.03 ± 0.33
e TrkIso < 3.0	97.63 ± 0.12	98.16 ± 0.11	96.58 ± 0.51
τ EVeto	92.88 ± 0.21	92.80 ± 0.22	92.10 ± 0.46
τN Prongs $= 1$	10.12 ± 0.25	7.591 ± 0.23	6.976 ± 0.65
τ EcalIso < 1	15.78 ± 0.95	17.08 ± 1.20	14.10 ± 1.57
τ TrackIso< 1	17.75 ± 2.51	15.06 ± 2.78	16.28 ± 3.93
$\cos\Delta\phi(e,\tau) < -0.95$	56.10 ± 7.75	40.00 ± 9.80	42.86 ± 8.98
$Q(e) * Q(\tau_{seed}) < 0$	60.87 ± 10.18	70.00 ± 114.4	33.33 ± 18.70
$\not\!\!E_{\text{T}} > 30$	21.43 ± 10.97	42.86 ± 18.7	0
p_ζ	66.67 ± 27.22	33.33 ± 27.2	0
b-Tagging	100.00 ± 0.00	100.00 ± 0.00	0

VII.6.6 Optimization

The values used for each selection criteria described above were obtained after optimization studies. The optimization was carried out by studying different cuts for each selection criteria and choosing the value that produces the lowest limit for the signal cross section (the cross section limit is explained in Chapter X). Figure 48 shows the cross section limit on the y axis versus different values of missing transverse energy.

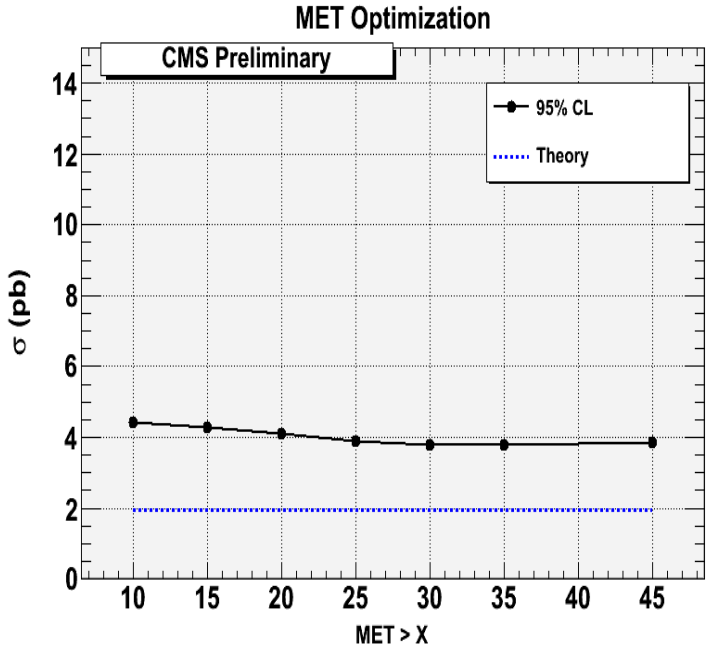
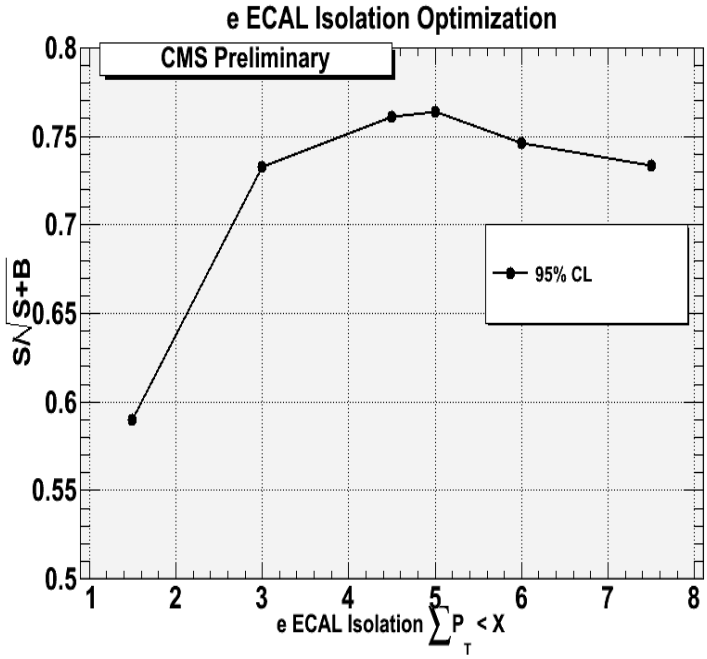


Fig. 48. Optimization for the missing transverse energy. The figure shows the cross section limit versus different values of missing transverse energy. The optimal point corresponds to the value that yields the lowest cross section limit.

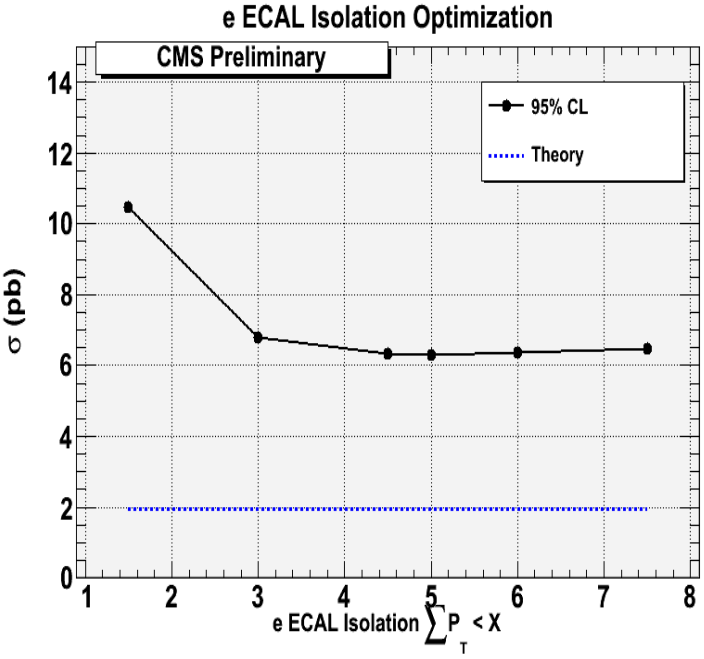
We also compared our results with a standard method used in particle physics that looks at the significance of simulated signal versus background. The idea is to obtain the highest signal to background ratio using the formula shown in Equation 53. The value of the selection to be optimized that yields that highest significance, is chosen at the optimal point.

$$S_g = \frac{S}{\sqrt{S+B}} \tag{53}$$

Figure 49 shows an example for optimization of the electron Ecal isolation obtained using the best significance and the lowest cross section limit. Notice that both methods yield the same optimal value.



(a)



(b)

Fig. 49. Example of optimization process for the electron Ecal isolation. (a) optimization done using signal to background significance, (b) optimization done by obtaining the best cross section limit. Notice that both methods yield the same optimal point. The blue line corresponds to the theoretical cross section value for a Z' with a mass of 500 GeV.

CHAPTER VIII

BACKGROUND ESTIMATION

In order to estimate the contribution of the different backgrounds that passed our selections, we use data driven methods wherever possible. The data driven technique consists of creating regions where a particular background has a dominant contribution with respect to other types of events. In order to create these regions, commonly referred to as control regions, we loosen or reverse some of the selection criteria to enhance the contribution of the background that we would like to estimate. Then, we use the collision data to measure the efficiency of the loosened selections in these control regions in order to extrapolate back to the signal region. The extrapolation is done by scaling down the number of data events in the control region by the value of the efficiencies of the loosened selections in the signal region.

The selection criteria used to create the enhanced regions should be uncorrelated to avoid biases and an incorrect efficiency measurement of the cuts which can lead to an over or underestimation of the background in the signal region. If two of the selection criteria used to create an enhanced region for a particular background are correlated, their efficiency is measured jointly.

If a data driven estimation is not possible, we create an enhanced region for the background with high purity and measure a scale factor between data and MC ($Sc_f = Data/BG$) to rescale the MC expectation in the signal region (see Equation 54).

$$N_{BG_i}^{signal} = N_{BG_i}^{MC\ signal} \times Sc_f \tag{54}$$

VIII.1 QCD Estimation

To enhance the contribution of QCD events we loosen and remove selections that mainly suppress this background. As explained in section VII.3, QCD events are characterized by a high multiplicity of soft tracks, which makes isolation a powerful

discriminant with respect to tau jets. By loosening the electron and tau track isolation we enhance significantly the QCD fake rate. Because QCD events come from fragmentation of quarks and gluons, the missing transverse energy is expected to be small. By removing selections related to the missing transverse energy, in addition to the electron and tau isolation for tracks, we expect to obtain a region with an enhanced statistics of QCD events.

Starting with the standard selections described in Section VII.6 we loosen and remove the following selection criteria to obtain a statistically significant sample of events:

- "Loose" electron track isolation : $0 < eTrkIso < 15$
- "Loose" τ track isolation : $0 < \tau TrkIso < 15$
- SS and OS : $Q(e) * Q(\tau_{seed}) < 0 \; || Q(e) * Q(\tau_{seed}) > 0$
- No $\not\!\!E_{\rm T}$
- No ζ

Figure 50 shows the electron track isolation distribution obtained from events passing the above selection criteria. A clean sample of QCD events can be obtained by applying an anti-isolation cut on the electron ($4 < eTrkIso < 15$). Figures [51- 54] show the distributions for the tau track isolation, ζ, missing transverse energy and reconstructed mass after applying the anti-track isolation selection for electrons. The number of events for data and MC in this QCD region, QCD Region 1, is shown in Table 16.

The difference in the yields for data and MC is due to preselections applied in the production of the QCD simulated events. The QCD sample was produced applying loose selections in isolation for tracks and photons. Nevertheless, there is a good agreement in the overall shape between the collision data and the simulation after applying an overall scale factor of 1.38. The scale factor comes from the ratio between the number of events for data and QCD, presented in Table 16. Even

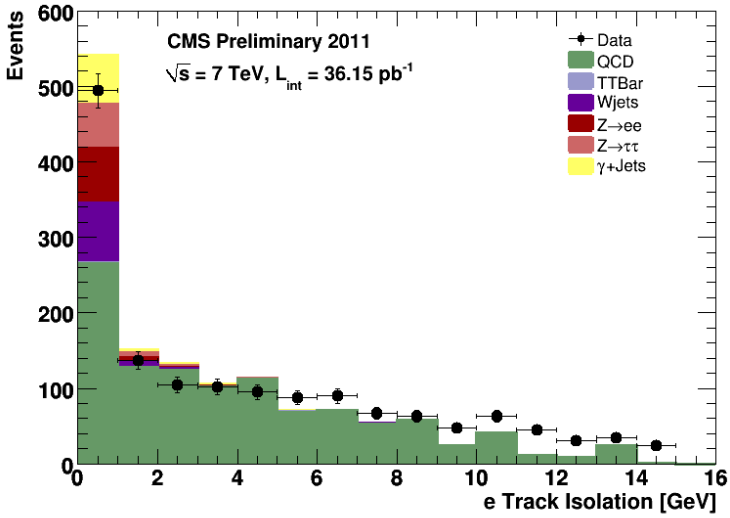


Fig. 50. Electron Track Isolation with selection criteria designed to enhance QCD jet events.

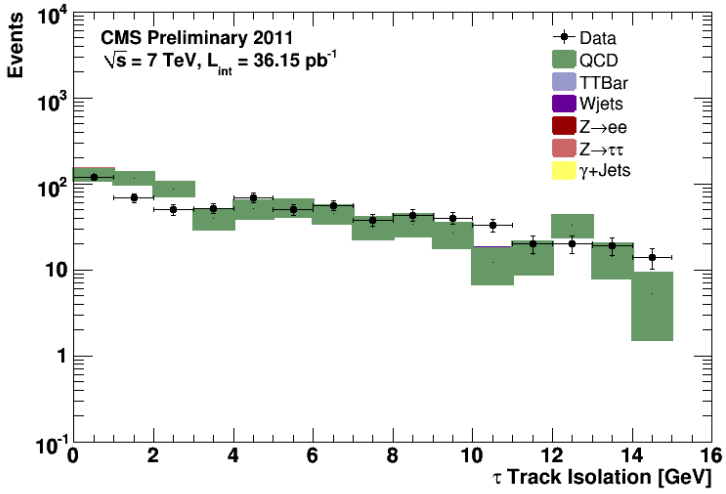


Fig. 51. Tau track isolation distribution for the $e\tau$ in the enhanced QCD control region 1.

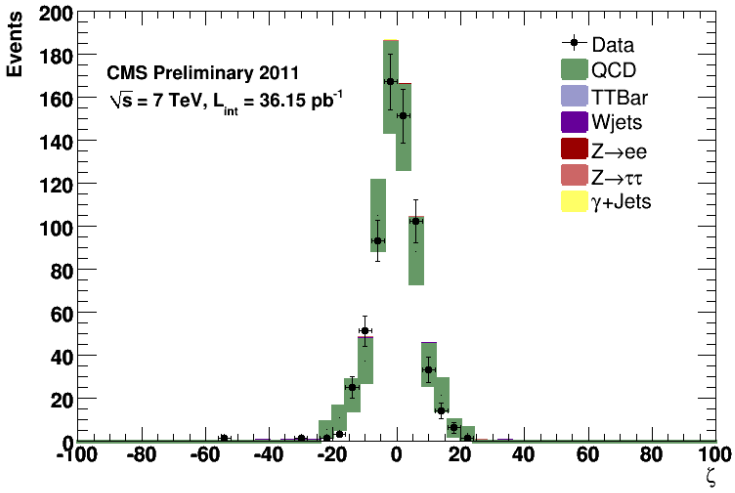
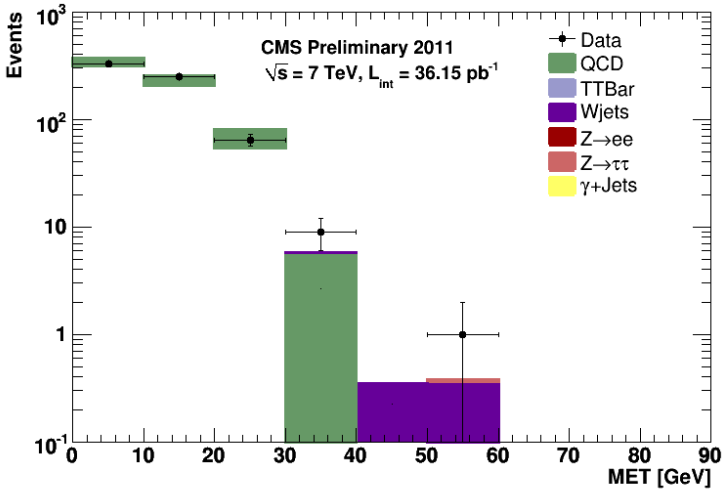
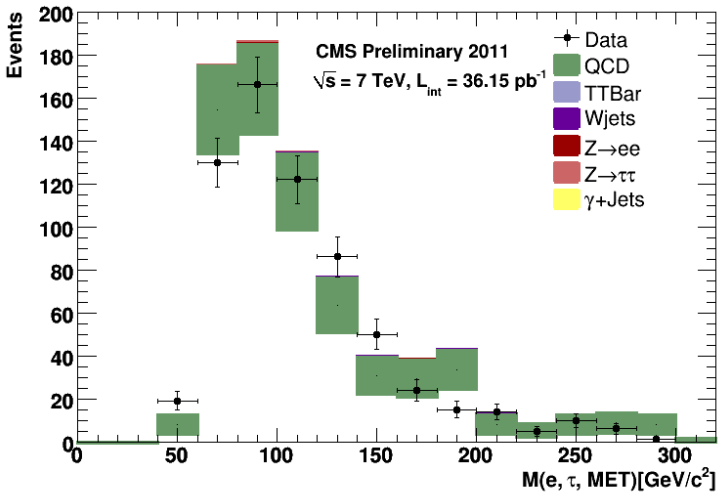


Fig. 52. p_ζ distributions for the $e\tau$ in the enhanced QCD control region 1.



(a)

Fig. 53. Missing transverse energy distribution for the $e\tau$ in the enhanced QCD control region 1.



(a)

Fig. 54. $M(e, \tau, \not{E}_{\mathrm{T}})$ distribution for the $e\tau$ in the enhanced QCD control region 1.

Table 16
Events in QCD control region 1 for data and MC.

Sample	Events
Data	648
QCD	469.53 ± 37.64
$t\bar{t}$	0.13 ± 0.02
$W + Jets$	1.61 ± 0.34
$Z \rightarrow \tau\tau$	1.77 ± 0.23
$Z \rightarrow ee$	0.84 ± 0.15
$\gamma + Jets$	0.74 ± 0.72
Total MC Events	474.62 ± 37.64
Purity	0.99
$Sc_f(Data/QCD_{mc})$	1.38 ± 0.12

though agreement in the shapes between data and the simulation is good, this is not necessary as the background estimation is carried out in a data driven way. This means that we use the data and not the simulation to measure efficiencies and extrapolate to the signal region. We also can observe that the contamination in the QCD control region 1 coming from other backgrounds is very minimal.

In order to measure the electron track isolation efficiency, we need to create a second clean sample of QCD events. This is done by applying loose selection criteria as described above, but using an anti-track isolation cut on the hadronic tau leg ($4 < \tau TrkIso < 15$) as shown in Figure 55.

Having the clean QCD regions and all the efficiencies of the cuts used to obtain them, we can determine the expected number of QCD events in the signal region. This number is calculated as shown in Equation 55.

$$N_{QCD}^{signal} = \underbrace{N_{QCD}^{pure} \varepsilon^{\tau TrkIso<1} \varepsilon^{\not{E}_T>30GeV,\ \zeta>-7} \varepsilon^{Q(e)*Q(\tau)<0}}_{QCD\ region\ 1} \underbrace{\frac{N^{eTrkIso<3.5}}{N^{4<eTrkIso<15}}}_{QCD\ region\ 2} \tag{55}$$

The efficiencies and the expected number of QCD events in the signal region are shown in Table 17.

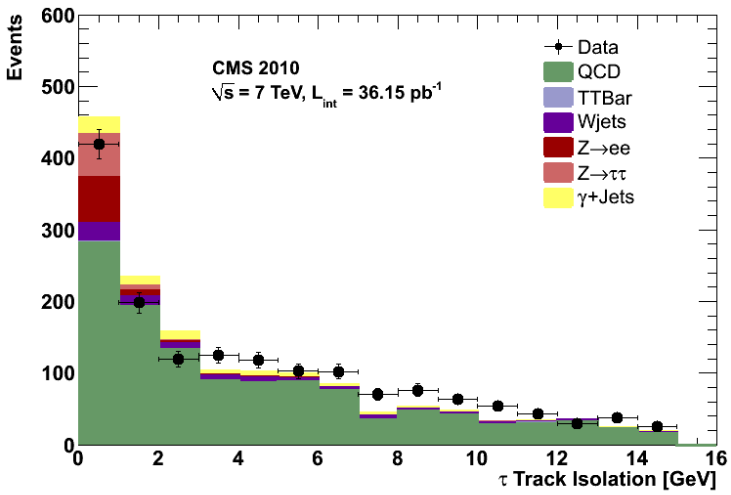


Fig. 55. Tau Track Isolation with selection criteria designed to enhance QCD jet events. By applying an anti-isolation cut, $4 < \tau TrkIso < 15$, we can obtain a second control region for QCD events in order to measure the electron track isolation efficiency.

Table 17
Efficiencies calculated from data in QCD control region.

Cut	Efficiency(%)
$\varepsilon^{\tau TrkIso<1}$	23.27 ± 1.61
$\varepsilon^{\not{E}_T>30GeV,\ \zeta>-7}$	0.93 ± 0.38
$\varepsilon^{Q(e)*Q(\tau)<0}$	52.70 ± 1.96
$\varepsilon^{eTrkIso}(N^{eTrkIso<3.5}/N^{4<eTrkIso<15})$	85.83 ± 1.82
Expected Number of Events	0.63 ± 0.26

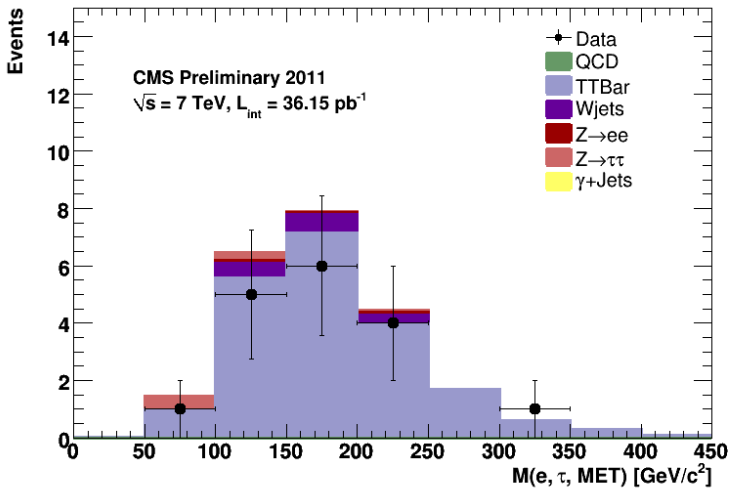
VIII.2 $t\bar{t}$ Extraction

The extraction of this background is performed in a data driven way. Due to the presence of the extra jets coming from the fragmentation of the b quarks in the topology of this background, the decay products that fake the electron and the hadronic tau are smeared with respect to the jet direction. This characteristic implies that $t\bar{t}$ events faking the $Z' \rightarrow \tau\tau \rightarrow e + \tau$ signature are not necessarily back to back in $\eta - \phi$ space. For this reason, the $cos\Delta\phi(e, \tau_h)$ and ζ cuts are powerful discriminants to suppress this background. Another selection criteria used to remove $t\bar{t}$ events, is to require zero jets tagged as b jets (Section VII.6).

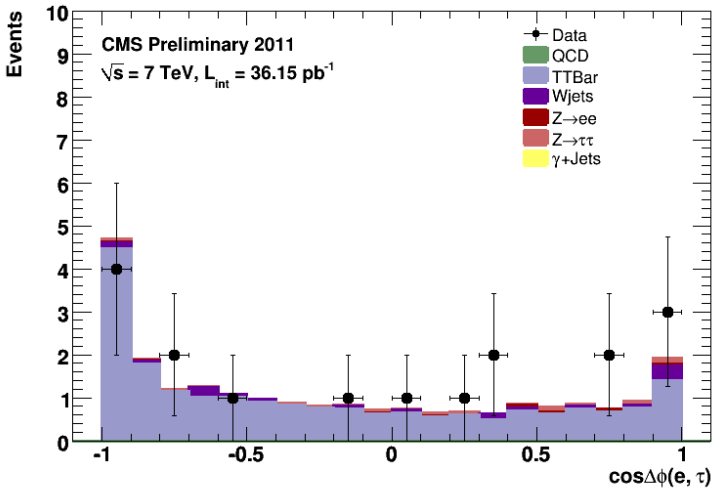
In order to create a region dominated by $t\bar{t}$ events we modify the following selections from our standard criteria:

- Remove requirement on ζ cut
- Remove requirement on $cos\Delta\phi(e, \tau_h)$
- Define a jet as $\Delta R(e/\tau, jet) > 0.5$ and require ≥ 1 jet tagged as a b-jet using the track counting high efficiency "low" discriminator (TCHEL)

We use the track counting high efficiency "loose" working point tagger, instead of the "medium" value used for the selections in the signal region, to enhance statistics in the $t\bar{t}$ control region. The loose working point use a relaxed selection criteria to tag a jet as a b-jet [30].



(a)



(b)

Fig. 56. (a) $M(e, \tau, \not\!\!E_{\mathrm{T}})$, and (b) $cos\Delta\phi(e, \tau_h)$ distributions for the $t\bar{t}$ enhanced region.

Figure 56 shows the mass and $cos\Delta\phi(e, \tau_h)$ distributions in this control region. Table 18 shows the number of events for data and MC.

In order to extrapolate to the signal region we use Equation 56:

Table 18
Events in $t\bar{t}$ control region for data and MC.

Sample	Events
Data	17
QCD	0
$t\bar{t}$	20.94 ± 0.32
$W + Jets$	1.40 ± 0.31
$Z \rightarrow \tau\tau$	0.87 ± 0.16
$Z \rightarrow ee$	0.31 ± 0.09
Total MC Events	23.52 ± 0.48
Purity	0.89
Sc_f	0.81 ± 0.20

$$N_{t\bar{t}}^{Signal} = N_{t\bar{t}}^{pure} \frac{P^{TCHEM}(0\ b\text{-jets})}{P^{TCHEL}(1\ b\text{-jet}) + P^{TCHEL}(2\ b\text{-jets})} \varepsilon^{cos\Delta\phi(\mu,\tau)} \varepsilon^{\zeta} \tag{56}$$

where $P^{TCHEM}(0\ b\text{-jets})$, $P^{TCHEL}(1\ b\text{-jet})$ and $P^{TCHEL}(2\ b\text{-jets})$ are the probabilities of having zero, one and two b-jets tagged with the TCHEM (zero b-jets) and TCHEL (one and two $b-jets$) algorithms.

The probability of tagging just one $b-jet$ is determined by the efficiency of the algorithm and it is defined as:

$$P(1\ b\text{-jets}) = (1 - \varepsilon_{b_1}) \times \varepsilon_{b_2} + (1 - \varepsilon_{b_2}) \times \varepsilon_{b_1} \tag{57}$$

where b_1 and b_2 label the two $b-jets$ coming from the $t\bar{t}$ decay. In Equation 57, ε_{b_1} represents the probability of the b-tagging algorithm to tag one $b-jet$. The index "1" is used to label one of the two candidates in the tau pair. Similarly, $(1 - \varepsilon_{b_i})$ represents the probability to not tag a $b-jet$ in the event. Since the efficiency to tag b_1 or b_2 is the same for each candidate in the tau pair, we can write equation 57 as $P(1\ b\text{-jets}) = 2 \times (1 - \varepsilon_b) \times \varepsilon_b$.

Using the same logic, we can easily show that the probability of tagging two b-jets is defined as:

$$P(2\ b\text{-jet}) = \varepsilon_b^2 \tag{58}$$

Equation 59 summarizes the equations to tag zero, one and two b-jets:

$$\begin{aligned} P(0\ b\text{-jets}) &= 1 - P(1\ b\text{-jets}) - P(2\ b\text{-jets}) \\ P(1\ b\text{-jets}) &= 2 \times (1 - \varepsilon_b) \times \varepsilon_b \\ P(2\ b\text{-jet}) &= \varepsilon_b^2 \end{aligned} \tag{59}$$

The selection efficiencies measured in the $t\bar{t}$ control region are shown in Table 19.

Table 19
Efficiencies calculated from data in the $t\bar{t}$ control region

Cut	Data %	MC %
$\varepsilon^{cos\Delta\phi(e,\tau)}$	11.80 ± 7.80	15.66 ± 0.56
ε^{ζ}	35.30 ± 11.60	49.40 ± 0.77
$\varepsilon^{bTagging}$(TCHEM)	45.50 ± 1.60	49.40 ± 0.30
$\varepsilon^{bTagging}$(TCHEL)	56.20 ± 2.00	63.6 ± 0.30
Probability to tag $\geq 1b - jets$ (TCHEL)	80.82 ± 4.30	86.75 ± 0.65
Probability to tag $0b - jets$ (TCHEM)	29.70 ± 2.50	25.60 ± 0.63
Expected Number of Events	0.26 ± 0.19	0.48 ± 0.02

VIII.3 $W + Jets$ Extraction

We extract the $W + jet(s)$ background in a data driven way. As explained in section VI.2, this background enters into our selection criteria when a W and an uncorrelated jet coming from a different process fake the leptonic (electron) or the hadronic tau. Because the W and the jet are uncorrelated, they are not necessarily back to back. For this reason the $cos\Delta\phi(e, \tau_h)$ and ζ selections are good discriminants to reduce this background. We can obtain a statistically significant sample of $W + Jets$ events by removing these two selections from our standard criteria. Unlike the

$t\bar{t}$ estimation, we keep the b-tagging selection requiring zero jets tagged as $b-jets$. Furthermore, to remove $Z \rightarrow \tau\tau$ events and achieve a better purity of $W + Jets$ events, we apply a cut in the transverse mass between the electron and the missing transverse energy.

The transverse mass is reconstructed using the components in the $x-y$ plane of the energy and the momentum of the electron and the missing transverse energy.

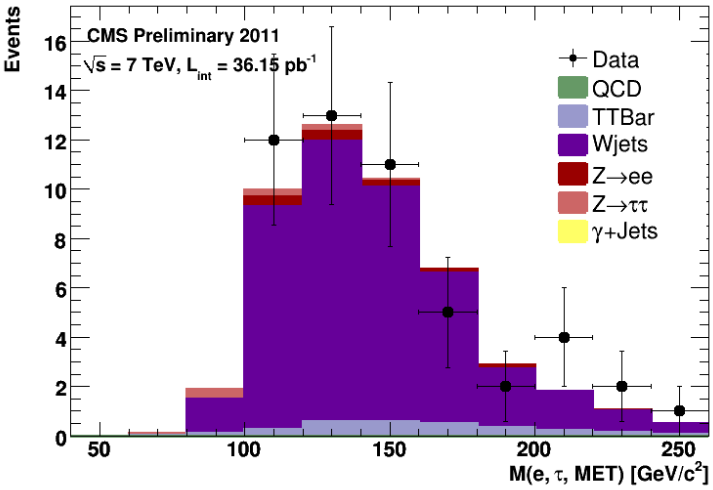
In order to create an enhanced region with these events we use the following selections:

- Begin with the selections outlined in section VII.6
- Remove the requirement on ζ
- Remove the requirement on $cos\Delta\phi(e, \tau_h)$
- Define a jet as $\Delta R(e/\tau, jet) > 0.5$ and require 0 jets tagged as b-jets using the track counting high efficiency "medium" discriminator (TCHEM)
- $50 < M_T(e, \not\!\!E_{\rm T}) < 100$ (removes contamination from $Z \rightarrow \tau\tau$ events)

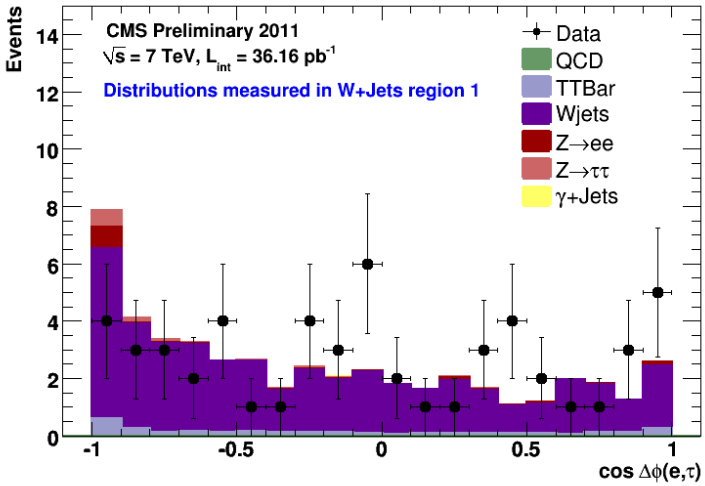
Figures 57 and 58 show the distributions for the mass, ζ and transverse mass between the electron and the missing transverse energy. Table 20 shows the measured efficiencies and expected number of events in the $W + jet(s)$ enhanced region.

In order to measure the efficiency of the $M_T(e, \not\!\!E_{\rm T})$ cut, we need to obtain a second control region of $W + Jets$ events. We can obtain it by applying all standard selections for the signal region, except the threshold in the transverse momentum for the hadronic tau. This is a consequence of the dominant soft momentum spectrum of jets and $W's$ that fake the hadronic tau.

The measured efficiencies and expected number of events in the signal region are presented in Table 21.

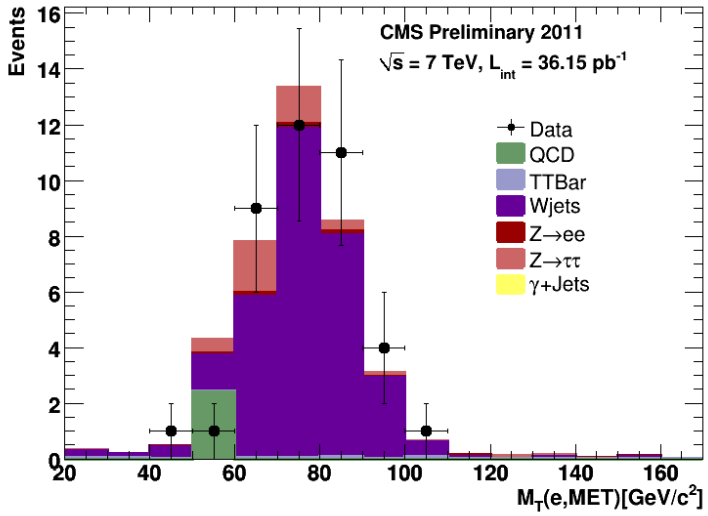


(a)



(b)

Fig. 57. Distributions in the $W + jets$ control region one. (a) $M(e, \tau, \not{E}_{\mathrm{T}})$, (b) ζ.



(a)

Fig. 58. Transverse mass distribution between the electron and the missing transverse energy in the $W + jets$ control region two ($M_T(e, \not{E}_T)$). After applying electron and tau identification selections the electron from the Z' decay is mostly faked by a real electron coming from the W. Notice that the transverse mass distribution peaks at around 80 GeV/c^2 in agreement with the mass of the W boson as expected.

Table 20
Events in $W + Jets$ control region for data and MC.

Sample	Events
Data	54
QCD	0
$t\bar{t}$	2.95 ± 0.12
$W + Jets$	43.7 ± 1.75
$Z \to \tau\tau$	1.13 ± 0.18
$Z \to ee$	1.4 ± 0.20
$\gamma + Jets$	0.02 ± 0.02
Total MC Events	49.21 ± 1.77
Purity	0.88
Sc_f	1.23 ± 0.17

Table 21
Efficiencies and expected number of events for the $W + Jets$ control region

Cut	Data %	MC %
$\varepsilon^{cos\Delta\phi(e,\tau),\zeta}$	1.85 ± 1.84	4.49 ± 3.13
$\varepsilon^{M_T(e,\not{E}_T)}$	95.50 ± 3.44	95.63 ± 3.64
Expected Number of Events	1.05 ± 1.04	2.05 ± 1.44

VIII.4 $Z \rightarrow e^+e^-$ Extraction

As explained in Section VI.2, this background enters our signal region because a prompt electron from the Z passes the electron selection criteria, and a bad quality electron fakes the hadronic tau jet. A bad quality electron, also known as a showering electron, is characterized by high bremsstrahlung radiation and a bad geometrical matching between the track position and the seed cluster position in the electromagnetic calorimeter, as explained in Section VII.4.

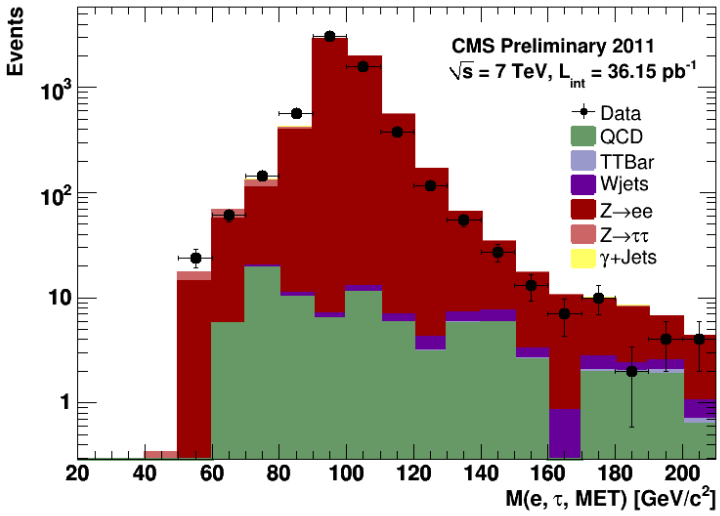
Around 96% of the showering electrons can be removed by applying the electron veto described in Section VII.6. Furthermore, since there are no associated neutrinos, a cut in the transverse missing energy reduces this background significantly.

Therefore, $Z \rightarrow e^+e^-$ can be enhanced with the following modifications to our selection criteria:

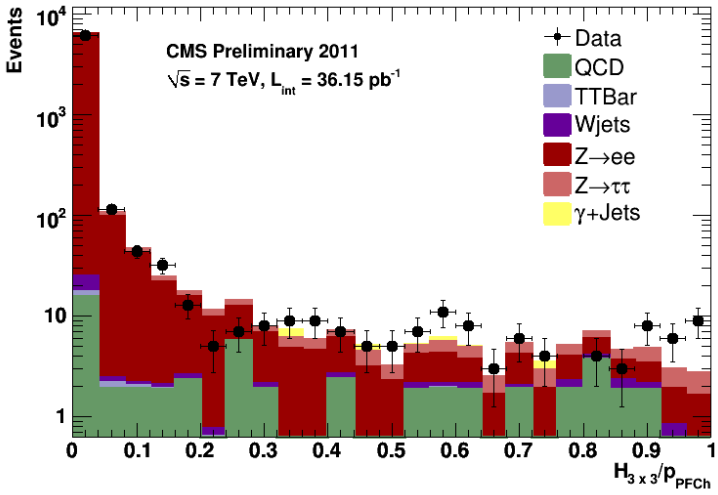
- Remove $H_{3\times3}/P_{\tau seed}$
- No $\not{E}_T$

The mass and the $H_{3\times3}/P_{\tau seed}$ distributions are shown on Figure 59 and the missing transverse energy distribution is shown in Figure 60. The number of events in this control region is shown in Table 22.

Since the tails of the $H_{3\times3}/P_{\tau seed}$ and the missing transverse energy distributions suffer from non-negligible contamination of QCD, $Z \rightarrow \tau\tau$ and $t\bar{t}$ events, we cannot measure their efficiencies from the control sample obtained with the above selection criteria. This prevents us from estimating the contribution of this background in a



(a)

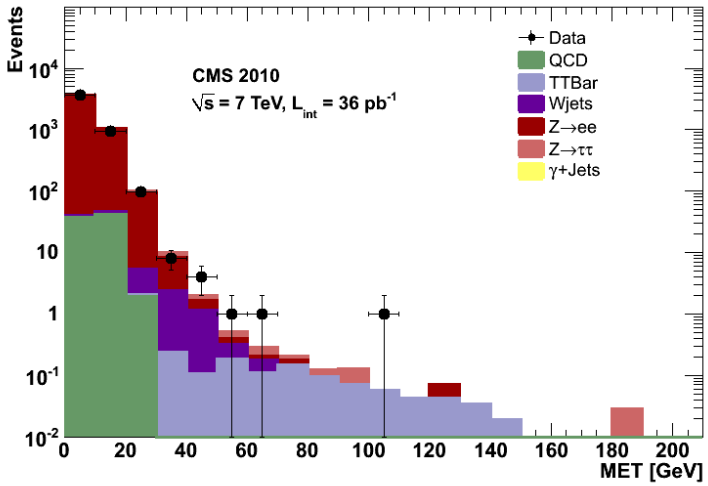


(b)

Fig. 59. Distributions for the $Z \rightarrow e^+e^-$ control region. (a) $M(e, \tau, \not{E}_{\mathrm{T}})$, (b) $H_{3\times3}/P_{\tau\ seed}$

Table 22
Events in $Z \to e^+e^-$ control region for data and MC.

Sample	Events
Data	4638
QCD	82.13 ± 17.30
$t\bar{t}$	1.55 ± 0.09
$W + Jets$	12.49 ± 0.94
$Z \to \tau\tau$	67.21 ± 1.39
$Z \to ee$	4665.00 ± 10.65
$\gamma + Jets$	8.38 ± 2.43
Total MC Events	4836.76 ± 20.53
Purity	0.96
Sc_f	0.99 ± 0.014



(a)

Fig. 60. Missing transverse energy distributions for the $Z \to e^+e^-$ control region.

data driven way. Instead, the expected contribution of this background in the signal region is obtained by using a data-MC scale factor:

$$N_{Z\to ee}^{signal} = N_{Z\to ee}^{MC} \times Sc_f = 0.75 \pm 0.15 \tag{60}$$

VIII.5 $Z \rightarrow \tau\tau$ Validation in Control Region

We do not employ a data-driven extraction method to estimate the $Z \rightarrow \tau\tau$ contribution in the signal region. Instead, we use the MC expectation for this background. Nevertheless, we show consistency between data and MC by creating an enhanced region with $Z \rightarrow \tau\tau$ events and subtracting the contamination with other backgrounds as described below.

In order to enhance the contribution of $Z \rightarrow \tau\tau$ events, we loosen the cut in the missing transverse energy from 30 to 5 GeV ($\not\!\!E_{\rm T} > 5$ GeV). The 5 GeV threshold was selected in order to reduce the contamination from $Z \rightarrow ee$ and QCD events as much as possible. To further suppress QCD events, we required a tighter electron isolation criteria ($eTrkIso < 1$ and $eECalIso < 1$).

We estimate the QCD contribution by extrapolating from the pure QCD control region 1 (described above) to the $Z \rightarrow \tau\tau$ enhanced region using Equation 61.

$$N_{QCD}^{Z\rightarrow\tau\tau} = \underbrace{N_{QCD}^{pure}\varepsilon^{\tau TrkIso}\varepsilon^{\not\!\!E_{\rm T}>5GeV,\ \zeta>-7}\varepsilon^{Q(e)*Q(\tau)<0}}_{QCD\ region\ 1}\ \underbrace{\varepsilon^{eTrkIso<1}\varepsilon^{eEcalIso<1}}_{QCD\ region\ 2} \tag{61}$$

$$\varepsilon^{eTrkIso<1} = \frac{N^{eTrkIso<1}}{N^{4<eTrkIso<15}}\ , \quad \varepsilon^{eEcalIso<1} = \frac{N^{eEcalIso<1}}{N^{eEcalIso<4.5}}$$

Table 23 shows the selection efficiencies used to extrapolate from the QCD enhanced region to the $Z \rightarrow \tau\tau$ enhanced region.

To further suppress the contamination from $Z \rightarrow ee$, a veto cut has been applied. Events where an $e\tau$ pair has an invariant mass $M(e,e)$ that lies within 3 σ of the nominal Z mass as reported in the PDG or pairs with an asymmetry in the transverse momentum ($p_{T\ asym} = p_T^{\tau} - p_T^{e} < 20$ GeV) are flagged as coming from $Z \rightarrow ee$ and rejected at the analysis level. This veto cut has an efficiency of 55.56 ± 9.56 % for $Z \rightarrow ee$ and 93.24 ± 2.92 % for $Z \rightarrow \tau\tau$ events.

To estimate the contribution of $Z \rightarrow ee$, $t\bar{t}$ and $W + Jets$ events, we use the background enhanced regions described in sections VIII.3 and VIII.4 to obtain scale

Table 23

Efficiencies calculated from data in the QCD control region one to extrapolate to $Z \rightarrow \tau\tau$ region .

Cut	Efficiency %
$\varepsilon^{\tau TrkIso}$	23.27 ± 1.61
$\varepsilon^{\not{E}_{\rm T}>5,\ \zeta>-7}$	66.05 ± 1.86
$\varepsilon^{Q(e)*Q(\tau)<0}$	52.70 ± 1.96
$\varepsilon^{eTrkIso<1}$	45.23 ± 2.60
$\varepsilon^{eEcalIso<1}$	10.26 ± 1.16
Expected Number of Events	2.44 ± 0.37

factors to be applied to the MC estimation in the $Z \rightarrow \tau\tau$ enhanced region (Equation 62). The scale factors for each background are shown in Tables 22, 18 and 20.

$$N^i_{Expected} = N^i_{MC} \times Sc^i_f \tag{62}$$

The number of events for data and all backgrounds are shown in Table 24.

Table 24

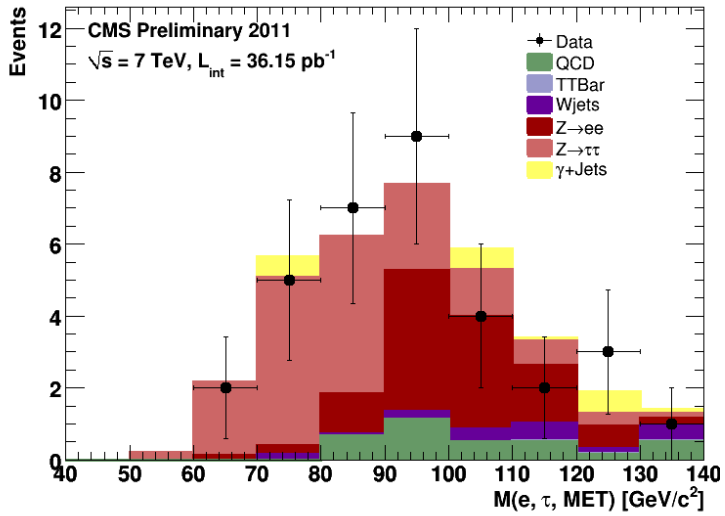
Events in the $Z \rightarrow \tau\tau$ enhanced region for data and MC. The QCD contribution has been calculated from data. Scale factors have been applied to $Z \rightarrow ee$, $W + Jets$ and $t\bar{t}$

Sample	Events
Data	39
QCD(data)	2.44 ± 0.37
$t\bar{t}$	0.14 ± 0.02
$W + Jets$	4.30 ± 0.64
$Z \rightarrow \tau\tau$	16.50 ± 0.69
$Z \rightarrow ee$	14.96 ± 0.64
$\gamma + Jets$	2.099 ± 1.21
Total MC Events	40.04 ± 1.70

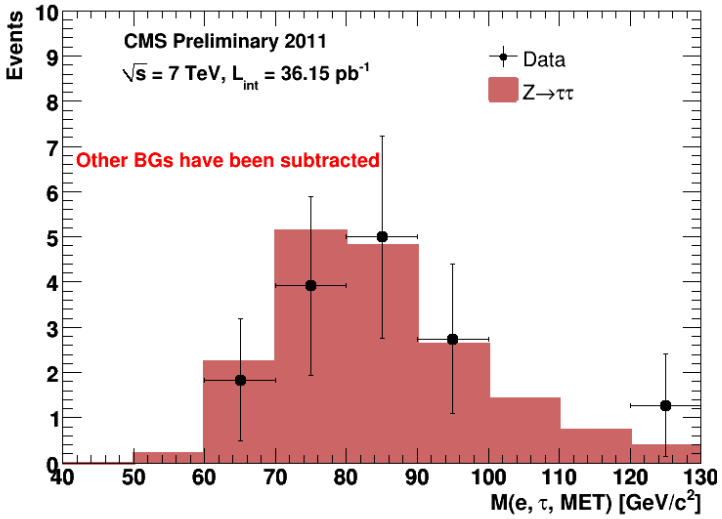
If we subtract from data the expected contribution of all backgrounds (Equation 63) we obtain an estimation of 15.46 ± 9.12 events, which is statistically consistent with the MC estimation for $Z \rightarrow \tau\tau$ (16.50 ± 0.69). Figure 61 shows the mass distribution before and after subtracting the contamination from other backgrounds. The $Z \rightarrow \tau\tau$ mass shape in MC is consistent with the mass shape obtained from the $Z \rightarrow \tau\tau$ enhanced region after subtracting the contamination from other backgrounds.

$$N^{Z\rightarrow\tau\tau} = N_{data} - N_{data}^{QCD} - N_{mc}^{Z\rightarrow ee} \times Sc_f^{Z\rightarrow ee} - N_{mc}^{t\bar{t}} \times Sc_f^{t\bar{t}} - N_{mc}^{W+Jets} \times Sc_f^{W+Jets} - N^{\gamma+Jets} \quad (63)$$

The expected number of $Z \rightarrow \tau\tau$ events in MC passing all the signal selection criteria is 2.15 ± 0.17.



(a)



(b)

Fig. 61. $M(e, \tau, \not{E}_{\mathrm{T}})$ for $Z \rightarrow \tau\tau$ enhanced region (a) before subtracting backgrounds, (b) after performing the subtraction bin by bin.

CHAPTER IX

SYSTEMATIC ERRORS

The main source of systematic uncertainty in our analysis comes from the estimation of the backgrounds in the signal region as result of the lack of statistics. There is a small effect introduced by the contamination with other backgrounds in the different control regions. In cases where a data driven estimation could not be performed, the systematic uncertainty is driven by the uncertainty in the expected number of events in the signal region for the particular background. Nevertheless, in most cases this contamination is small as shown in Chapter VIII and is negligible with respect to the uncertainty introduced by the lack of statistics.

After the background estimation, the tau identification uncertainty is the second most significant source of error in our analysis. As explained in Section VII.3, we identify taus using a shrinking cone algorithm and an elliptical isolation for photons. This identification algorithm is subject to different systematic effects:

- Efficiency of finding a track [31] associated with each charged hadron
- Efficiency of finding a track convoluted with the probability to have a tau with three tracks that are collinear such that it is identified as a tau candidate with one track. This can happen for taus with a very high p_T.
- Probability for charged pions or neutral pions to fall out of the signal cones.
- Probability for tracks and/or photons coming from underlying events (or pile up) to fall into the isolation cone.
- Probability for tracks coming from underlying events to fall in to the isolation cone spoiling the one or three prong requirement.

We are not affected by the systematic effect from three prong taus falling out of the isolation region because we select one prong hadronic taus for this analysis.

Nevertheless, we are still subject to three prong taus being identified as a one prong tau. This is a consequence of badly reconstructed tracks being associated to the one charged hadron. This effect is measured by using simulated taus (MC) from known processes such as $Z \rightarrow \tau\tau$ and counting how many reconstructed three prong taus, matched to generator level taus in the simulation, were identified as a one prong tau. The matching is carried out by finding the closest reconstructed candidate in η space to the generated tau. The result found by the tau physics object group is that 0.74% of taus with three prongs will be indetified as a one prong tau. This result includes taus with very high transverse momentum, which makes it a small effect.

The probability for tracks and/or photons coming from underlying events or pile-up to fall in to the isolation region is the same for taus, electrons and muons if the isolation cones are the same. For this reason, this effect can be measured by applying the tag and probe technique [28]. We use $Z \rightarrow \mu\mu$ events for this which have lower fake rates and are easier to identify than electrons and taus.

The uncertainty of finding a track associated with each charged hadron is measured using the ratio of neutral charm meson decays to two or four charged particles [31]. It was found that around 4% of all the reconstructed charged pions did not have an associated track. This value was included as part of the total systematic uncertainty for tau identification.

Although the described systematic effects for tau identification give an overall combined value of less than 5%, we use a conservative value of 7% recomended by the tau physics object group and used for other tau analyzes in CMS.

All of the other sources of systematic uncertainty were included or calculated using recommended values from the CMS physics object group.

The uncertainty introduced by the lack of knowledge of the parton distribution functions (PDF), was determined by comparing CTEQ6.6L PDF with the default PDF and variations within the CTEQ6.6 family of parametrizations.

Systematic effects related to the tau or electron energy scale were calculated by using $Z' \rightarrow \tau\tau$ and $Z \rightarrow \tau\tau$ simulated samples and shifting the energy or momentum

distributions by a constant factor using the generator level information, as shown in the formula below:

$$p_T^{new} = p_T^{gen} \times \sigma \tag{64}$$

Where σ is a constant factor recommended by the CMS. The value of p_T^{new} was propagated to all the selection criteria that depended on it. The value of the systematic effect introduced by the electron and tau energy scale is calculated as the difference in the final limit using the shifted and non-shifted momentum.

Similarly, the systematic effects introduced by the tau or electron momentum resolution were calculated using the generated and reconstructed information for each object:

$$p_T^{new} = p_T^{gen} + (p_T^{reco} - p_T^{gen}) \times \sigma \tag{65}$$

Where once again σ is a constant factor recommended by the CMS physics object group. The effect of this systematic uncertainty was calculated as the difference in the final limit using the "smeared" momentum resolution and non-smeared momentum resolution.

In Table 25, we present the complete list of systematic uncertainties considered in this analysis.

Table 25
Systematics for MC and Data

Source of Systematic	$e - \tau$
Luminosity	4%
Electron Trigger	0.39%
Electron ID	1.37%
Tau ID	7.0%
Parton Distribution Functions	3.96%
Initial State Radiation	2.14%
Final State Radiation	1.7%
Tau Energy Scale (3%)	2.1%
Electron Energy Scale (1%)	1.8%
Tau Energy Resolution	Negligible
Electron Energy Resolution	Negligible
Background Estimation	45%

CHAPTER X

STATISTICAL ANALYSIS

Experiments that are performed with n independent trials that yield a true or false outcome with a probability of success p, follow a binomial distribution $B(n,p)$. The binomial probability of obtaining l successes in a particular experiment given n independent trials is given by:

$$f(l,n,p) = \frac{n!}{n!(n-l)!}p^l(1-p^{n-l}) \tag{66}$$

The Poisson distribution is a special case of a binomial probability where the number of independent experiments is large and the probability of success is small:

$$\begin{aligned} &\lim_{n\to\infty} f(l,n,p) = \lim_{n\to\infty} \tfrac{n!}{n!(n-l)!}p^l(1-p^{n-l}) \\ &\lim_{n\to\infty} f(l,n,p) = \tfrac{\mu^l \exp^{-\mu}}{l!} \end{aligned} \tag{67}$$

where $\mu = np$. In Equation 67 "l" represents the number of observed events and μ the number of expected events defined as:

$$\mu_i = L_i \sigma_{sig} \varepsilon_i + b_i \tag{68}$$

where L_i is the integrated luminosity, σ_{sig} is the signal cross section, ε_i is the cumulative efficiency after applying all the selection criteria and b_i is the number of background events estimated using the methods outlined in Chapter VIII. In a simple counting experiment the above equations are sufficient to quantify the signal significance; however, this analysis uses a binned likelihood to achieve greater sensitivity. We find the tau-tau mass distribution provides the best separation between signal and background. Therefore, the sub-index "i" represents the expected and observed number of event in each bin of the mass distribution. The total likelihood using all bins is given by Equation 69.

$$\mathcal{L}(\varepsilon_1, \varepsilon_2 ..... \varepsilon_n) = \prod_i^{Nbins} \mathcal{L}(\mu_i, l_i) = \prod_i^{Nbins} \frac{\mu_i^{l_i} \exp^{-\mu_i}}{l_i!} \tag{69}$$

Since this analysis is one of several possible searches in final states with taus, the result must be combined with other tau decay channels to determine the overall assessment of a possible discovery or exclusion. The joint likelihood was calculated as shown in Equation 70.

$$\mathcal{L}_{final} = \mathcal{L}_{e\tau} \times \mathcal{L}_{\mu\tau} \times \mathcal{L}_{e\mu} \times \mathcal{L}_{\tau\tau} \tag{70}$$

The final likelihood is used to obtain the limit in the signal cross section. This limit is set using a 95% confidence level estimate. A confidence level sets the reliability of the experimental measurement established by the authors of the analysis, implying if an experiment is performed several times with the same physical conditions, a similar outcome will be obtained only 5% of the time. To obtain this confidence level, we integrate the likelihood as function of the signal cross section:

$$\frac{\int_0^{\sigma_{95}} \mathcal{L}(\sigma) d\sigma}{\int_0^{\infty} \mathcal{L}(\sigma) d\sigma} = 0.95 \tag{71}$$

We included systematic effects in our limits as nuisance parameters. In statistics, a nuisance parameter is a quantity that has no direct relation with the statistical sample but that needs to be introduced to account for possible changes that might affect the results of interest. A good example of a nuisance parameter is the variance of a normal distribution, which is introduced to measure the spread of the results with respect to the statistical mean. In our case we introduced each nuisance parameter as a log-normal probability density function for normalizations and Gaussian for mass spectrum uncertainties, with the mean of the distributions located at the nominal value of the each accounted systematic. The log-normal distribution was used in order to obtain positive values evenly distributed in a continuous spectrum. This allowed us to study the impact of a particular systematic on our results after smearing its value.

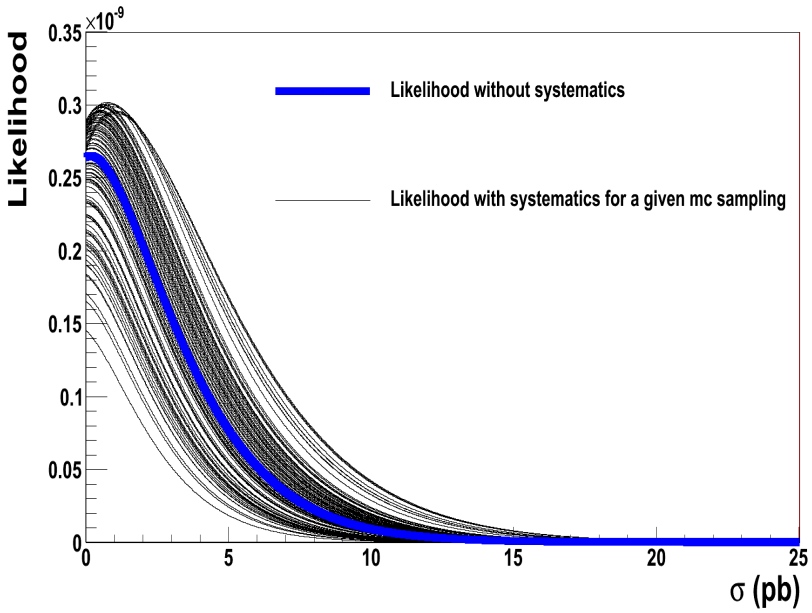


Fig. 62. Poisson likelihood. The distribution shows the likelihood without any smearing or nuisance parameters included, as well as several likelihood distributions that represent the effect of the nuisance parameters. The distributions were made using the expected number of events for the backgrounds in the $\mu\tau$ channel. The backgrounds were calculated applying data driven techniques.

Figure 62 shows the default likelihood without smearing or nuisance parameters included, as well as several likelihood distributions that represent the effect of the nuisance parameters.

Since we set a combined limit, there are correlated and uncorrelated errors across channels. For example, the systematic error introduced due to the imprecise measurement of the hadronic tau energy scale will be correlated across channels that search for one or two hadronic taus in their decay topology. The correlated errors considered for the $e\tau$ channel are:

- Tau momentum resolution (correlated with $\mu\tau$ and $\tau\tau$ channels)
- Tau energy scale (correlated with the $\mu\tau$ and $\tau\tau$ channels)
- Tau identification (correlated with the $\mu\tau$ and $\tau\tau$ channels)

- Electron momentum resolution (correlated with the $e\mu$ channel)
- Electron energy scale (correlated with the $e\mu$ channel)
- Particle Density Functions (correlated for all channels)
- Initial State Radiation (correlated for all channels)
- Final State Radiation (correlated for all channels)
- Luminosity (correlated for all channels)

The uncorrelated errors are:

- Background estimation
- Trigger errors

Equation 72 shows the expected number of events after systematic effects have been introduced:

$$\mu_i^{'} = (1 + g_L)L_i\sigma_{sig}(1 + f_{\varepsilon_i})(1 + g_\varepsilon)\varepsilon_i + (1 + f_{b_i})(1 + g_b)b_i \tag{72}$$

In Equation 72 g and f represent the correlated and uncorrelated factors.

In order to test the statistical framework and estimate the sensitivity of our selection criteria, we performed preliminary studies using simulated background only as input data (pseudo-data), using a Poisson based random event generator. Figure 63 shows an example of a MC based sensitivity study where pseudo-data was generated using samples from background only distributions.

We can have different scenarios depending on how the events in data are distributed after applying all our selection criteria. If data is located at the low mass of the spectrum in a consistent way with the expected backgrounds, our limit will have a low value according to the Poisson likelihood distribution. Nevertheless, if few events in data fluctuate towards the high mass region, this will result in high value in our limit, which can prevent us from setting a limit of exclusion or claiming a discovery if the statistics are low. Figure 64(a) shows an example (pseudo-experiment)

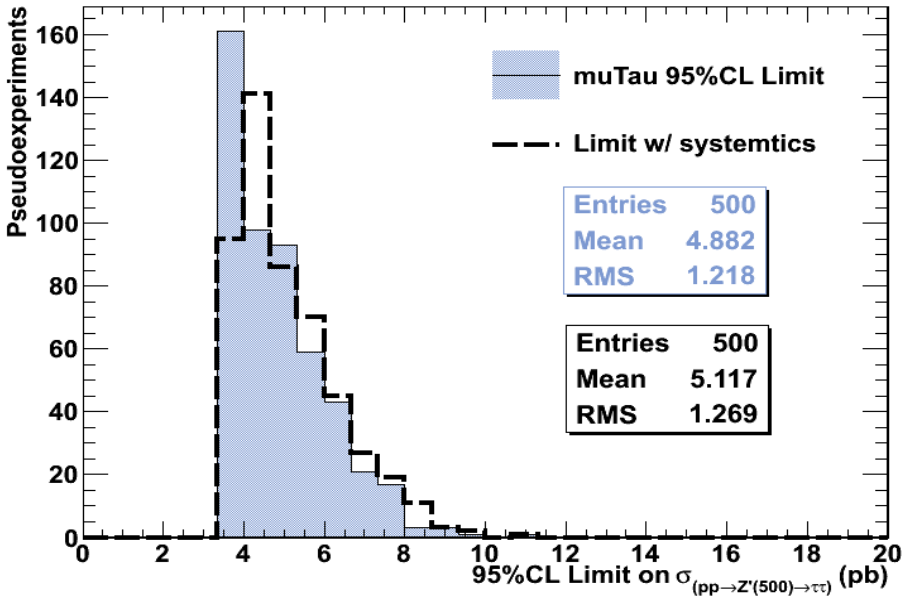


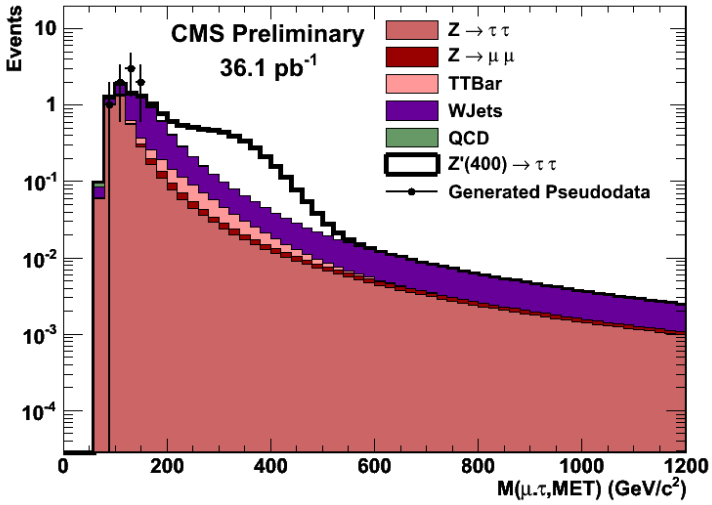
Fig. 63. 95% C.L. limits for a sample of pseudo experiments. The studies were performed using the $\mu\tau$ channels as example.

where the pseudo-data has mostly fallen in the low mass region, which results in an expected cross section of $\sigma_{95} = 4.4\ pb$. Figure 64(b) shows an example where data has fluctuated to the high mass region, thus resulting in a expected cross section vale of $\sigma_{95} = 12.8\ pb$ ("unlucky"). These studies allowed us to analyze fluctuations in our results in order to set the 95% C.L limit.

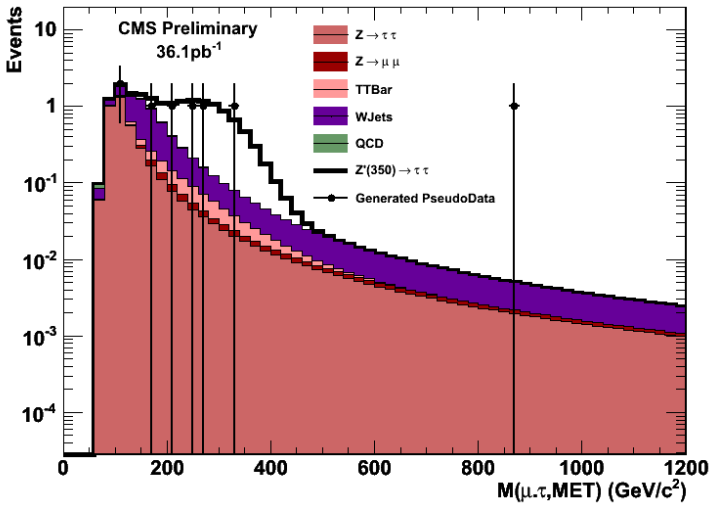
An imprecise knowledge of the mass shapes used to calculate the binned likelihood, can result in a bad estimate of the limit. For this reason, we used a "morphing" procedure in order to introduce variations in the mass shapes (smearing) to see the effect on the final limit. The smeared mass (D_i^j) templates were used to obtain new limits to compare them with the values given by the non-smeared mass shapes (D_i^{def}). The expression for the modified likelihood integral is given by:

$$N^{-1}\sum_{j=1}^{N}\mathcal{L}(\epsilon_1+\alpha_1^j\delta\epsilon_1,..,\epsilon_n+{\alpha_n}^j\delta\epsilon_n,D_1^{def}+\alpha_1^j\delta D_1^j,...,D_n^{def}+\alpha_n^j\delta D_n^j) \qquad (73)$$

where $\delta\mathcal{D}_i^k = \mathcal{D}_i^k - \mathcal{D}_i^{def}$ is the difference between the default and the deviated shape for the k^{th} systematic effect.



(a)



(b)

Fig. 64. (a) Example pseudo-experiment (with $Z'(400)$) resulting in $\sigma_{95} = 4.4\ pb$ (b) Example pseudo-experiment (with $Z'(350)$) resulting in $\sigma_{95} = 12.8\ pb$. These studies allowed us to analyze fluctuations in our results in order to set the 95% C.L limit. The studies were performed using the $\mu\tau$ channels as example.

CHAPTER XI

RESULTS

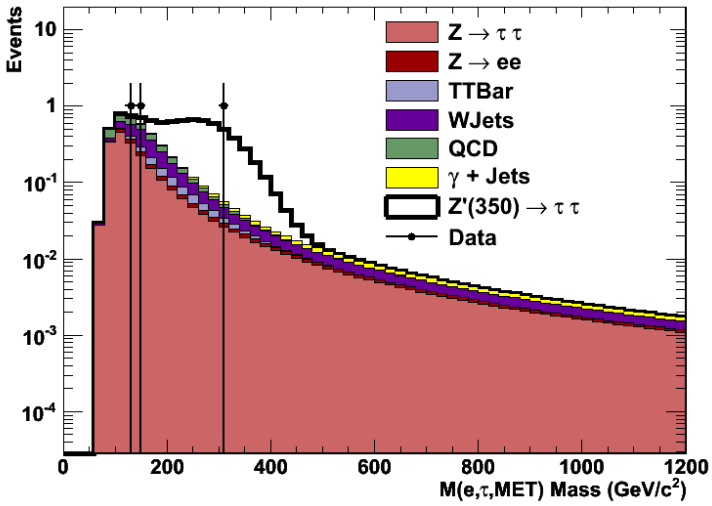
There is no evidence of a new mass resonance in the data analyzed. The number of events in the mass spectrum obtained from data agrees with the expected number of events in MC. The events are located in the low mass region.

The dominant backgrounds for the $e\tau$ channel are $W + jets$ and Drell-Yan processes. Table 26 shows the number of expected and observed events in the signal region, while Figure 65 shows the mass distributions in the signal region for backgrounds and simulated Z' masses of 350 and 700 GeV respectively. The masses were reconstructed using the visible products for the electron and the tau and the missing transverse energy.

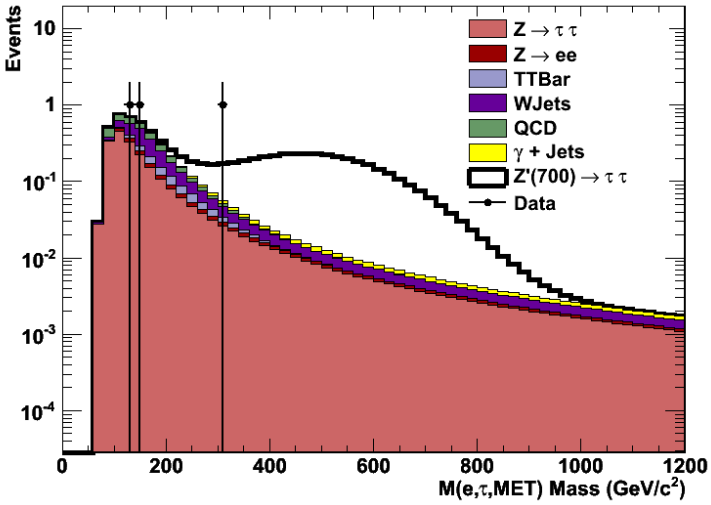
Due to the lack of MC statistics in our MC samples, we used fits to the mass shapes in order to obtain smooth tails. To obtain the fits, as shown in Figure 65, we loosened some of the selection criteria in order to increase the statistics of each particular background. The loosened selections were chosen such that the mass shapes were not biased. The shapes were scaled down to the expected number of events in the signal region. The fits are shown in Figures 66 to 68.

Table 26
$e\tau$ Events in the signal region.

Sample	Events
QCD	0.63 ± 0.26
$t\bar{t}$	0.26 ± 0.19
$W + Jets$	1.05 ± 1.04
$Z \rightarrow \tau\tau$	2.15 ± 0.26
$Z \rightarrow ee$	0.19 ± 0.05
$\gamma + Jets$	0.12 ± 0.12
Total	4.40 ± 1.14
Observed	3

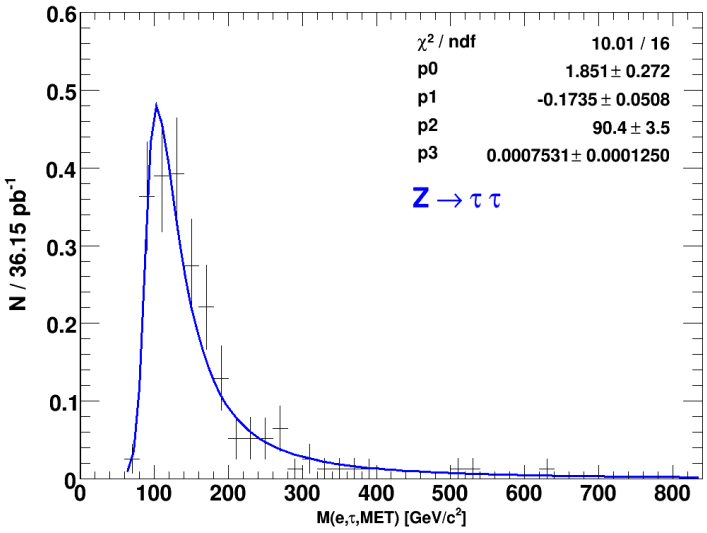


(a)

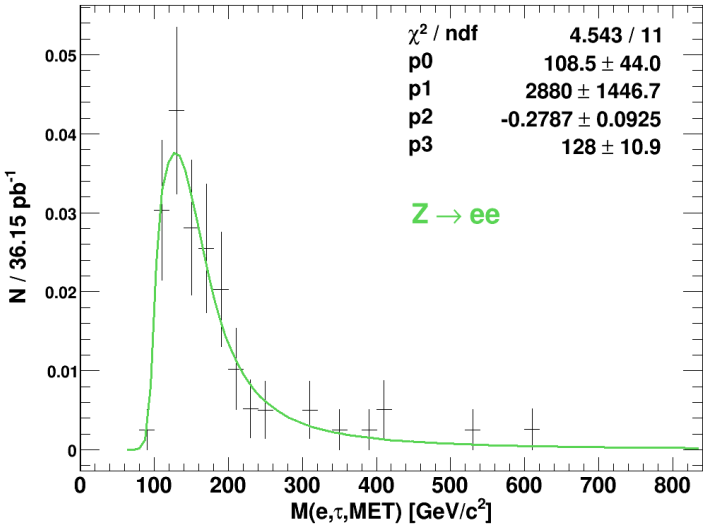


(b)

Fig. 65. $M(e, \tau, \not{E}_{\mathrm{T}})$ for $Z^{'} \rightarrow \tau\tau$ masses of (a) 350 GeV and (b) 700 GeV.

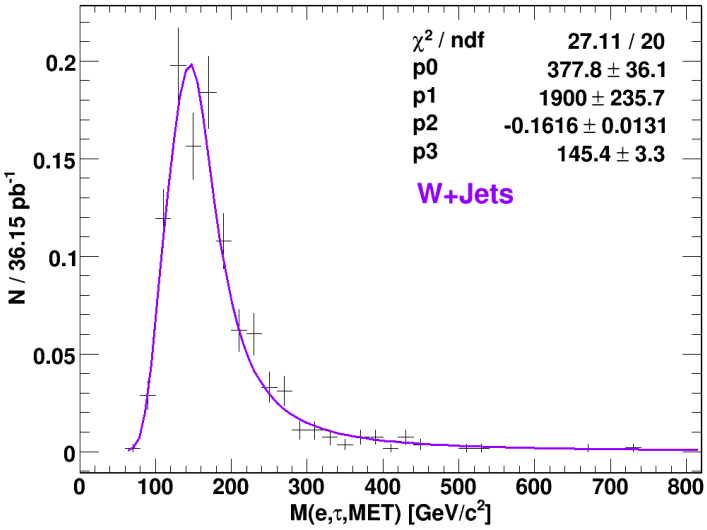


(a)

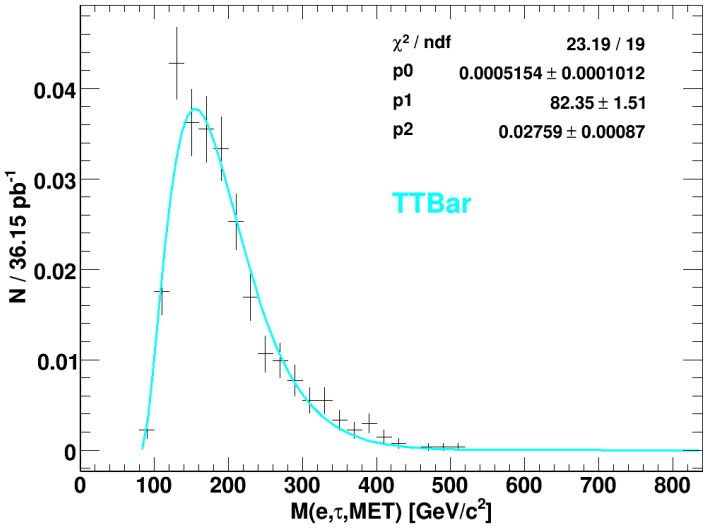


(b)

Fig. 66. Fits to the mass shapes ($M(e, \tau, \not\!\!E_{\mathrm{T}})$) of different backgrounds. (a) $Z \to \tau\tau$, (b) $Z \to ee$.

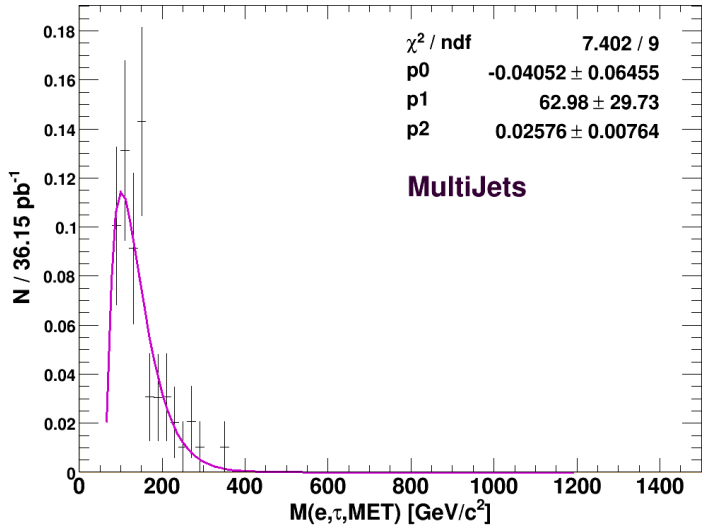


(a)

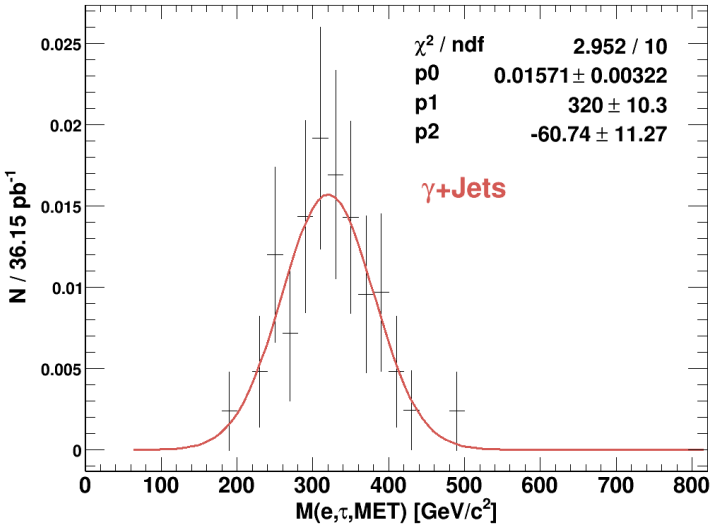


(b)

Fig. 67. Fits to the mass shapes ($M(e, \tau, \not\!\!E_{\mathrm{T}})$) of different backgrounds. (a) $W + Jets$, (b) $t\bar{t}$.



(a)



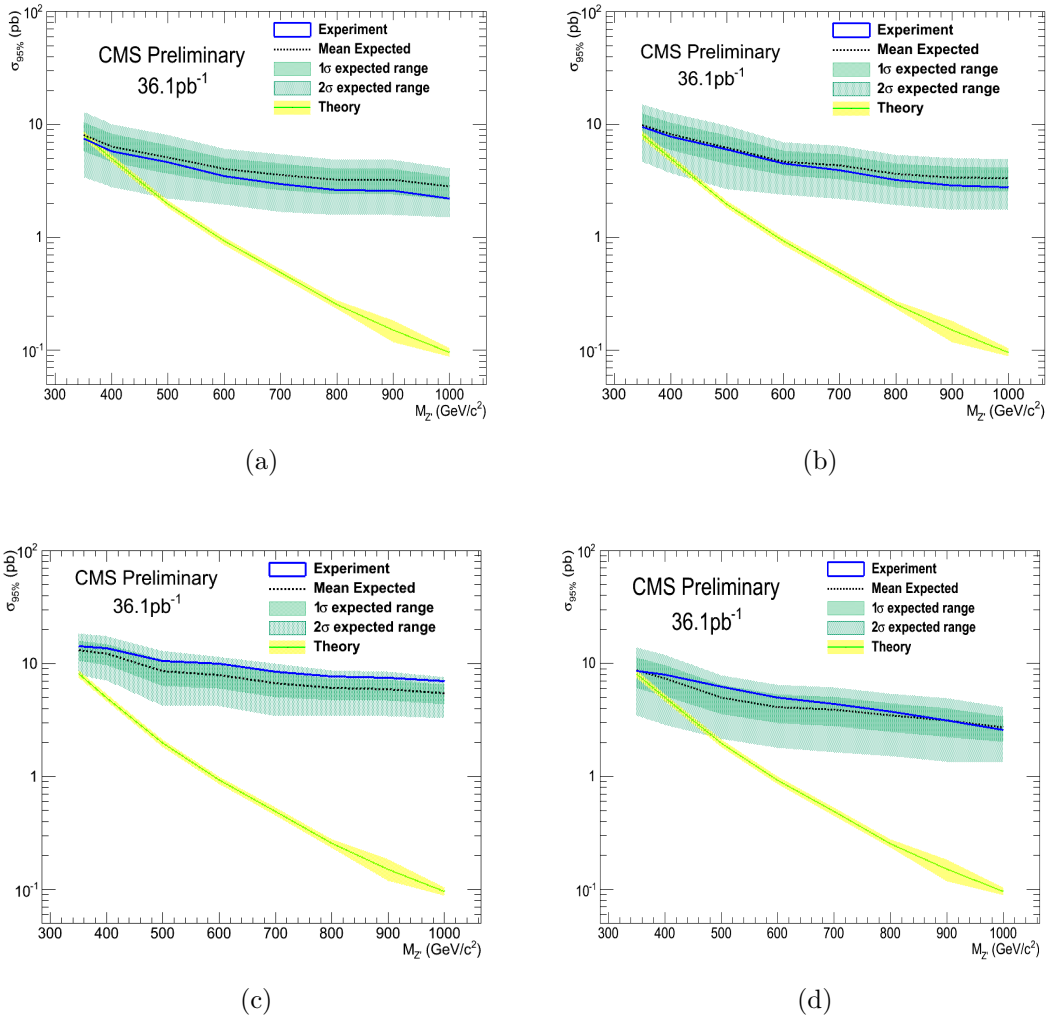
(b)

Fig. 68. Fits to the mass shapes ($M(e, \tau, \not{E}_{\mathrm{T}})$) of different backgrounds. (a) Multijets (QCD), (b) $\gamma + Jets$.

Since there is no evidence of a new mass resonance, we set an upper level limit on the cross section as a function of different Z' masses. The limit is set using a 95% confidence level as explained in Chapter X. The experimental and theoretical limits as a function of different Z' masses for the different channels in the combined limit are shown in Figure 69.

The joint limit is shown in Figure 70. We exclude the Z' a mass up to 468 GeV, which exceeds the sensitivity achieved by the CDF experiment [5] that excluded the $Z' \rightarrow \tau\tau$ up to 399 GeV.

The bands on the expected limits represent the 1σ deviation obtained using a large sample of pseudo-experiments where the pseudo-data are obtained from background only distributions using a Poisson based random event generator.



(a) (b)

(c) (d)

Fig. 69. 95% CL upper limits on the cross-section for (a) $\mu\tau$, (b) $e\tau$, (c) $e\mu$, (d) $\tau\tau$ as a function of the mass

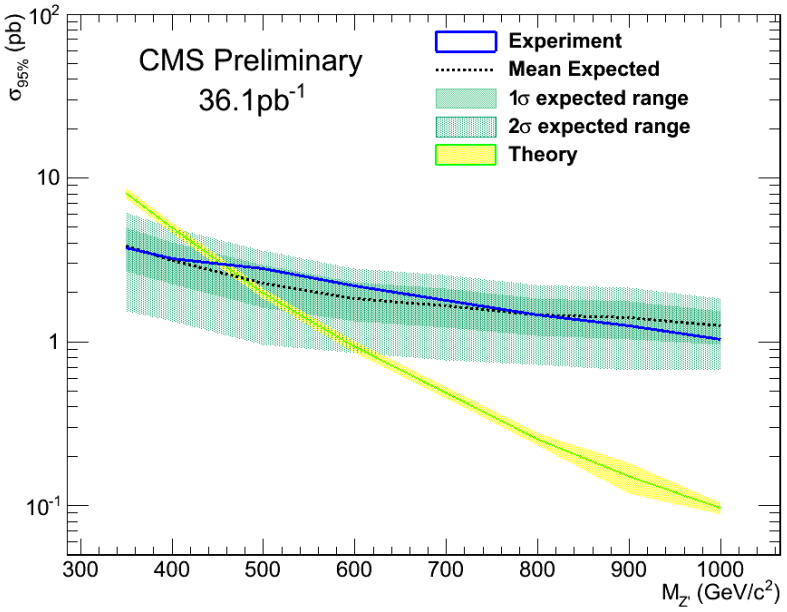


Fig. 70. 95% CL upper limits on the cross-section for join limit $= (\mu\tau) + (e\tau) + (e\mu) + (\tau\tau)$ as a function of the Z' mass

.

CHAPTER XII

IMPROVEMENTS FOR THE ANALYSIS AND EXTENDED WORK

As it was mentioned in Chapter IX, the main source of systematic uncertainty comes from the estimation of the backgrounds in the signal region as result of the lack of statistics. During the data taking period from March until August 2011, the LHC has collected an integrated luminosity of 2 fb^{-1}. This is around 55 times more statistics than for the 2010 data taking. I have been actively involved in the analysis of the collected data in 2011.

The selection criteria used for the 2011 analysis is the same than for the 2010 analysis but the hadronic tau identification methodology changed in order to use the recommended selections from the tau physics object group for the 2011 data. We have maintained most of our background estimation techniques, except for few modifications due to the new tau identification criteria. Due to the new trigger selections in 2011 we were able to change our control region to validate the selection criteria for tau identification using $Z \rightarrow \tau\tau$ events. In 2010, due to the trigger, we had to use a threshold of 25 GeV for the transverse momentum of the electron which significantly reduced $Z \rightarrow \tau\tau$ events since electrons from this background have an average a momentum of 15 GeV. For the 2011 run we are using a trigger that allows us to apply a threshold of 20 GeV for electrons. I have been mentoring a P.hD student from Texas A&M University who will be responsible for the analysis in future iterations. The analysis with 2011 data is expected to be approved by the end of August 2011 and will be submitted for publication in a physics journal.

XII.1 Supersymmetry Analysis in Final States with Taus

I have been actively involved in an analysis searching for signatures of supersymmetry in final states with hadronic taus. The production squarks ($\tilde{q}$) and gluinos ($\tilde{g}$) are expected to be the main source to discover supersymmetry at the LHC. Some decay topologies involving $\tilde{q}$ and $\tilde{g}$ can produce neutral particles known as neutralinos

which can decay to sypersymmetric taus ($\tilde{\tau}$) and tau leptons. A sketch of this decay process is shown in Figure 71. Since $\tilde{q}$ and $\tilde{g}$ are expected to be highly energetic, we search for events with at least two jets with high transverse momentum. The presence of neutralinos in the decay topology implies large missing transverse energy in the events. Furthermore, we search for two tau leptons decaying hadronically, coming from the neutralinos one ($\tilde{\chi_1^0}$) and two ($\tilde{\chi_2^0}$). As mentioned in Section V.7.1, in some models supersymmetry is more likely to be discovered in final states with taus. I have made many significant contributions to the analysis in the validation of the tau identification, studies of selection criteria to reduce backgrounds, trigger studies and estimation of different backgrounds. The analysis has been pre-approved by CMS and final approval is expected on August 15^{th} 2011.

XII.2 Other Contributions to the Experiment

A summary of other contributions to the CMS experiment is presented in Appendix

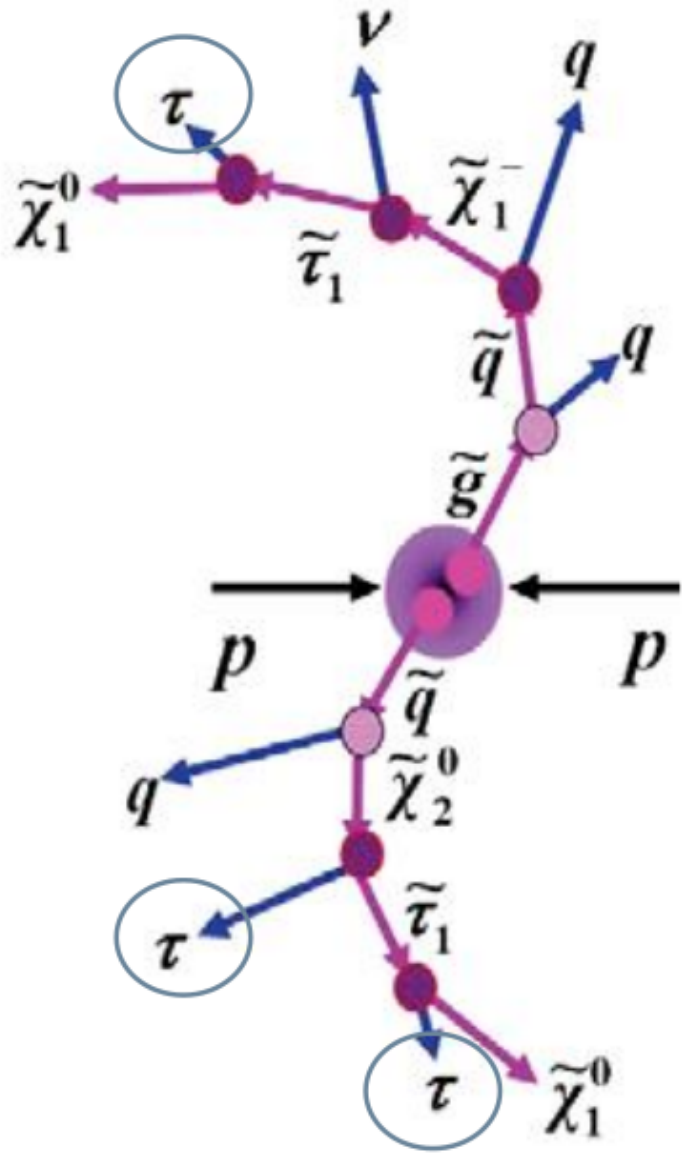


Fig. 71. Sketch of production and decay of supresymmetric particles producing taus in the final state.

APPENDIX I

OTHER CONTRIBUTIONS TO THE CMS EXPERIMENT

A summary of several other contributions to the CMS experiment is presented below.

I.1 Hardware Contributions for the Pixel Detector

In the fall and winter of 2007, I spent 3 months at CERN helping with the testing of the pixel detector. I developed a software interface for the readout of the pixel Front End Driver (FED) with the S-Link card. The S-Link allows data to be read out of the pixel FED at rates exceeding 400MB/s bypassing the VME bus. The interface used a finite state machine system to execute different processes according to user needs and allowed communication with the software for the detector control (PixelSupervisor). The software helped calibrations to occur quickly, and was the only fast readout available for pixel detector testing until the spring of 2008.

In cooperation with the PSI Physics institute of Switzerland, I helped with the installation and testing of the 32 FED's and 32 S-link cards for the Barrel Pixel Detector (Bpix).

During the early testing at the CERN Tracker Integration Facility (TIF), we were concerned about the noise and changes in the detector performance after it was shipped to CERN. I studied and compared all the results for noise obtained both at CERN and Fermilab. The study showed high consistency in the results.

With the collaboration of a technician from the INFI of Italy, I installed one of the two VME creates for the Fpix. I was in charge of the firmware update for the FEDs and FECs (Front End Controllers) at the TIF and was responsible for the installation of a set of power cables and optical fibers for one of the half service cylinders for the Fpix at the TIF.

During the Summer of 2008, I spent three months at CERN. I developed a routine using CMSSW (CMS Software) to retrieve private words coming in the data packet from the FED. The private words allow us to establish the progress during

calibrations and determine if we got the requested number of triggers (there is a pathology in the calibration delivery method that sometimes gives an unpredictable number of triggers). The routine retrieves the private words and stores them in a new data collection.

The Pixel Detector has a total of 40 FED's for its readout, 32 for the Bpix and 8 for the Fpix. Each pixel FED has 36 channels and not all of them are connected to a detector module. I studied the simulated channel occupancy in the FED. For instance, if a particular set of collisions produced more than 192 hits in a channel, that channel will produce an "overflow" which could cause a data loss, and, depending of the data packet length and structure, a loss of synchronization. I developed a simulation using CMSSW to verify and extend earlier work at PSI to determine channel occupancy. The results showed consistency with the PSI results. The data were simulated considering expected bunch crossings of protons per collision of 17.

In collaboration with a technician from the INFN of Italy, I helped in the installation of a new system of power supplies at the Tracker Integration Facility (TIF) at CERN. This system is now used as a test stand for the forward pixel detector's power supplies. I also helped with a several modifications carried out in the cooling lines of the pixel detector. I contributed with the data taking during the 2010 run as a DQM and pixel on-line shifter at P5.

I.2 Tau Trigger

I worked on the CMS tau trigger in collaboration with University of Wisconsin. This work consisted of determining the efficiency of the tau level 2 (L2) trigger with data and simulated data. Tau particles have some specific signatures that distinguish them from other particles. Low multiplicity is a nice characteristic of hadronic decays of taus, with 85% of the hadronic component decaying into states with a single charged pion. Thus, in order to identify taus that decay hadronically, we can look for a jet-like cluster in the calorimeters with a small number of charged tracks associated with it. Tau jets with high transverse momentum are very narrow,

which helps to discriminate them from sources of background such as QCD jets. In order to be reasonably confident that the trigger was performing well, we studied backgrounds that mimic taus in minimum bias data. Such fakes are available offline from the Particle Flow analysis. The Particle Flow (PF) technique uses energy deposits in the calorimeters and tracking information in order to reconstruct jets with good resolution. In order to translate the information from fake PF-taus into what is expected for real PF-taus, we use the Monte Carlo by comparing the behavior of fake PF-taus to "real" PF-taus from a known resonance, such as $Z \rightarrow \tau^+\tau^-$, for a variety of trigger variables. The level 1 (L1) trigger pre-selects tau jet candidates based on calorimetry energy thresholds and isolations. With more information available at the Level 2 (L2) trigger, we can refine the L1 criteria and take better advantage of the expected narrowness of the tau jet candidates at high transverse energies (E_t). Efficiency studies showed that the trigger reconstruction efficiency below 40 GeV is low which makes it difficult to discriminate real taus from fakes. I carried out studies to correct the difference in the low E_t region for fake and real taus. This correction is based on the parametrization of the reconstruction efficiency at low E_t of pure tau jets and fakes with MC in order to get correction factors that allow us to make a more precise identification of real tau jets with real data. I also conducted studies with the data delivered by the LHC. My work was based on the efficiency estimation for the L2 tau jet reconstruction, L2 tau jet energy resolution and efficiency of the L2 isolation algorithms.

I.3 EJTERM Contribution

CMS has organized a series of workshops meant to bring its physicists up to speed on the tools, software, and methods needed to contribute to physics analysis of CMS data. The EJTERM was organized at Fermilab with this aim. The Vanderbilt group was asked by the organizers of EJTERM to put together an exercise on Tau analysis. I wrote a basic code to analyze the $Z \rightarrow \tau^+\tau^-$ with one tau decaying hadronically and the other to a muon and associated neutrinos. I then assembled a set of exercises

for the participants to complete, with the aim of teaching them the basic physics and analysis tools used for tau identification.

I.4 Pixel Conditions Data Base Management

For four months, I was in charge of the conditions database for the CMS pixel detector. This database contains information related to the hardware, calibrations, geometry and other important information for the pixel detector. This information must be updated every time there is a change or a new set of run conditions in the detector. I created a web page that provides detailed instructions for database management.

REFERENCES

[1] G. Kane. Modern Elementary Particle Physics: The Fundamental Particles and Forces?. Preseus Publishing, 1993.

[2] *Unravelling an Extra Neutral Gauge Boson at the LHC Using Third Generation Fermions*, R. Diener et al., aXiv:hep-ph/1006.2845v1, 2010.

[3] *The Physics of Heavy Z-prime Gauge Bosons*, P. Langacker, Rev. Mod. Phys 81, 2009.

[4] *Search for Heavy Resonances Decaying to Tau Pairs in CMS*, J. Cumalat, N. Dhingra, A. Florez, S. Gennai, A. Gurrola, M. Hildreth,W. Johns, T. Kamon, E. Luiggi, K. Mazumdar, A. Safonov, J. Singh, I. Suarez, and N. Valls, CMS Analysis Note AN-11-007. *Search for Heavy Resonances Decaying to Tau Pairs in CMS*, J. Cumalat, N. Dhingra, A. Florez, S. Gennai, A. Gurrola, M. Hildreth,W. Johns, T. Kamon, E. Luiggi, K. Mazumdar, A. Safonov, J. Singh, I. Suarez, and N. Valls, CMS Analysis Note AN-11-007.

[5] *Search for New Physics Using High-Mass Tau Pairs from 1.96 Tev ppbar Collisions*, D. Acosta et at., Phys. Rev. Lett. 95, 2005.

[6] http://pdg.lbl.gov

[7] *Introduction to Electroweak Symmetry Breaking*, S. Dawson, http://arxiv.org/abs/hep-ph/9901280v1, 1999.

[8] Amitabha Lahiri, Palash B. Pal. A First Book of Quantum Field Theory. Alpha Science, 2007.

[9] *Low Energy Phenomenology of Superstring Inspired E(6) Models*, Hewett, J.L. and Rizzo, T.G., Phys. Rept. 183, 1989.

[10] *The CERN Large Hadron Collider: Accelerator and Experiments*, Volume 1, 2008.

[11] *The CERN Large Hadron Collider: Accelerator and Experiments*, Volume 2, 2008.

[12] W.-M. Yao et al., [Particle Data Group], Journal of Physics G 33, 1 (2006), and 2007 partial update for 2008 edition.

[13] *Determining the Dark Matter Relic Density in the Minimal Supergravity Stau-Neutralino Coannhiliation Region at the Large Hadron Collider*, R. Arnowitt, B. Dutta, A. Gurrola, T. Kamon, A. Krislock, and D. Toback, Phys. Rev. Lett. 100, 231802 (2008).

[14] *Missing Transverse Energy Performance in Minimum-Bias and Jet Events from Proton-Proton Collisions at* $\sqrt{s} = 7$ *TeV*, CMS Collaboration, CMS Physics Analysis Summary CMS-PAS-JME-10-004, http://cms-physics.web.cern.ch/cms-physics/public/JME-10-004-pas.pdf.

[15] *Measurement of the DrellYan Cross Section in pp Collisions at* $\sqrt{s} = 7$ *TeV*, CMS Collaboration, arXiv:submit/0293118, 2011.

[16] http://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html

[17] *CMS missing transverse energy performance in photon+X events for pp collisions at 7 TeV*, N. Akchurin, J. Damgov, S.W. Lee, E. Yazgan, CMS Analysis Note AN-10-432.

[18] *The MC@NLO 3.3 Event Generator*, S. Frixione, B. R. Webber, arXiv:hep-ph/0612272v1, 2006.

[19] *Evidence for Anomalous Lepton Production in* e^+ - e^- *Annihilation*, Per1 et al., 1975.

[20] T. Sj ostrand and others, JHEP 2006 026, 2006.

[21] http://madgraph.hep.uiuc.edu

[22] Was, Z. and others, Nucl. Phys. B, Proc. Suppl. 98, 2001.

[23] http://geant4.cern.ch

[24] *Parton fragmentation and string dynamics.*, B. Andersson, G. Gustafson, G. Ingleman, and T. Sjorstrand, Physics Reports, 97(2-3), 1983.

[25] *Tau reconstruction and identification with particle-flow techniques using the CMS detector at LHC*, Pioppi, M. and others, Nucl. Phys. B, Proc. Suppl. 189 (2009) 311 - 316.

[26] *Electron Reconstruction in CMS*, Baffioni. S et al., Eur. Phys. J. C 49, 1099 - 1116, 2007

[27] *Study of tau reconstruction algorithms using pp collisions data collected at* $\sqrt{s} =$ 7 *TeV*, CMS Collaboration, CMS Physics Analysis Summary PFT-10-004.

[28] *Electron and muon efficiency measurements in 2010 proton-proton dataset*, CMS Collaboration, CMS Physics Analysis Summary, CMS AN2010/464, 2010.

[29] *Search for High-Mass Tau Pairs in 1.96-TeV* $p\bar{p}$ *Collisions*, thesis by Z. Wan, unpublished, Rutgers University, 2005 (FERMILAB-THESIS-2005-13).

[30] *Commisioning of the b–Jet Identification with pp Collisions at* $\sqrt{s} = 7$ *TeV*, CMS Collaboration, CMS Physics Analysis Summary, http://cms-physics.web.cern.ch/cms-physics/public/BTV-10-001-pas.pdf , 2010.

[31] *Measurement of Tracking Efficiency*, CMS Collaboration, CMS Physics Analysis Summary TRK-10-002, 2010.