SEARCH FOR HEAVY RESONANCES DECAYING TO TAUS IN 7 TEV PROTON-PROTON COLLISIONS AT THE LARGE HADRON COLLIDER

A Dissertation

by

ALFREDO GURROLA III

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

August 2011

Major Subject: Physics

SEARCH FOR HEAVY RESONANCES DECAYING TO TAUS IN 7 TEV PROTON-PROTON COLLISIONS AT THE LARGE HADRON COLLIDER

A Dissertation

by

ALFREDO GURROLA III

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Teruki Kamon
Committee Members,	Alexei Safonov
	Richard Arnowitt
	Stephen Fulling
Head of Department,	Ed Fry

August 2011

Major Subject: Physics

ABSTRACT

Search for Heavy Resonances Decaying to Taus in 7 TeV Proton-Proton Collisions at the Large Hadron Collider.

(August 2011)

Alfredo Gurrola III, B.S., Texas A&M University

Chair of Advisory Committee: Dr. Teruki Kamon

Over the last few decades, the Standard Model (SM) of particle physics has been used as a means of understanding the world around us. However, there is an increasing amount of data that suggests the SM of particle physics only describes nature up to energies of the electroweak scale. Extensions to the SM have been developed as a means of explaining experimental observation. If these extensions are indeed the correct mathematical descriptions of nature, the Large Hadron Collider (LHC), located at the European Center for Nuclear Research (CERN) near Geneva, Switzerland, is expected to produce new and exciting physics signatures that can shed light on the evolution of our universe since the early hypothesized Big Bang. Of particular interest are models that may lead to events with highly energetic tau lepton pairs. In this dissertation, focus is placed on a possible search for new heavy gauge bosons decaying to highly energetic tau pairs using a data sample corresponding to an integrated luminosity of 36 pb^{-1} of proton-proton collisions at $\sqrt{s} = 7$ TeV collected with the CMS detector at the CERN LHC. The number of observed events in the data is in good agreement with the predictions for SM background processes. In the context of the Sequential SM, a Z' with mass less than 468 GeV/c^2 is excluded at 95% credibility level, exceeding the sensitivity by the Tevatron experiments at the Fermi National Accelerator Laboratory.

For as long as I can remember, I knew I wanted to be a scientist. I knew I wanted to be a doctor. As a kid, I used to sit outside at night with my parents and stare at the stars, stare at the blackness and vast emptiness of outer-space and wonder how we came to be. I used to wonder what was floating around in outer-space and where did it end. What were stars and planets and how did they get there? My parents used to tell me that someday I would help answer those questions. They used to tell me that they had a feeling that I was meant to be a scientist. That I was meant to do great things. For as long as I can recall, I've worked hard to be the best person, the best student, the best physicist that I could be. I wanted to meet my own goals of being a doctor. I wanted to make the people in my life proud. This dissertation is dedicated to my entire family. I hope I have made everyone proud and I hope I have been a good role model. This dissertation is especially dedicated to my wife, Leslie Michelle Gurrola, and my parents, Alfredo Gurrola II and Dolores Gurrola. My parents always told me to fall in love with someone that made me a better person and shared similar goals in life. I found the love of my life when I was 17 years old, and she has been by my side ever since, accompanying me all over the world and making great sacrifices to help me accomplish my goal of being a doctor and making me a better person along the way. My wife has watched me struggle though the doctoral program, has watched me have sleepless nights studying or working on research projects. She has been patient having the confidence that I would persevere. Leslie, I love you and thank you for always being there right by my side. Mom and Dad: Since I was a young kid looking up at the stars and constantly asking "why?", you told me to always try to be the best that I could be. You taught me the words I live by: "What matters most is how smart you are, how strong you are, and how big your heart is." You taught me to compete, to be strong, to work hard, and to be humble. More importantly, you taught me to persevere, and to get back up when I was down. Words can't express how grateful I am. I love both of you very much and I hope I have made you proud.

ACKNOWLEDGMENTS

When I first arrived as a graduate student at Texas A&M University, I had no idea what type of physics I wanted to pursue or how a graduate student would go about finding an advisor to help guide their dissertation. I decided to talk to the academic advisor at the time, Teruki Kamon. I had absolutely no idea that fate would decide my path that day. I was lucky enough, and blessed, that Teruki Kamon decided that day to collaborate and guide me on a project that now defines my research. Dr. Kamon has been the perfect advisor, mentor, and source of wisdom. To Dr. Kamon I want to say thank you for all the guidance. Thank you for teaching me how to be a physicist. Thank you for understanding the struggles of a graduate student, both academically and personally. Thank you for the great physics talks, and I look forward to the continued collaboration in the future. There's no doubt that Dr. Kamon will continue to be a mentor and a source of wisdom, academically and personally.

Dr. Kamon allowed me to work with some incredible people that have really shaped the way I think about physics. I would like to thank Professor Bhaskar Dutta for his guidance. My first research project was under the supervision of Dr. Dutta and Dr. Kamon. Professor Dutta gave me great insight into the connection between theory and experiment and the growing field of phenomenology. Dr. Dutta constantly encouraged me to find the connection between the design of an experimental study and the theoretical motivations behind them. Since I was a kid, I have always been interested in the makeup of the universe and the possibilities that are floating around in outer-space. Dr. Dutta has given me the opportunity to explore and study the possibilities of phenomena well beyond our current understanding of physics. I have never met anyone that enjoys physics as much as Professor Bhaskar

Dutta and I hope that I can carry the same enthusiasm in my future endeavors.

While working with Dr. Dutta and Dr. Kamon on searches for new physics at the LHC, I naturally fell in love with physics searches involving tau leptons. Professor Alexei Safonov, as one of the leading tau experts in the world, gave me the opportunity to work with him and has played a large part in guiding my tau research projects over the past five years. Several people have played a large part in my development as a physicist, however, Alexei Safonov was one of those that always pushed to get the maximum out of me. I expressed my desire to be the best that I could possibly be, and Professor Safonov always pushed to get the most out of me. He always provided additional opportunities for me to work with other students, professors, and experts around the world. To Professor Alexei Safonov, I want to say thank you for the guidance. Thank you for the opportunities you have provided me. Thank you for always pushing me to think outside the box.

One of the most enjoyable experiences for me as a graduate student was the physics conversations I had with Professor Richard Arnowitt. It's not everyday that someone has an opportunity to work with such a well respected physicist. From Professor Richard Arnowitt I learned how to make the more important physics connections to fields or studies outside the scope of a particular research project. Very often these are the most difficult connections to make, and Professor Arnowitt helped guide me in making those physics connections.

I would also like to thank Professor Dave Toback for the many great physics conversations and his guidance on the experimental aspects of an analysis. As a young physicist who would think of the underlying physics in terms of a purely mathematical perspective, Dr. Toback helped guide my understanding of the many subtleties

involved with high energy physics at hadron colliders.

I was given an extraordinary opportunity to work at the European Center for Nuclear Research (CERN) near Geneva, Switzerland. Being far away from Texas A&M, Professor Mayda Velasco from Northwestern University guided my work on the the anomalous behavior of the CMS detector and its relevance to the measurement of important physics quantities. More importantly, Professor Mayda Velasco was a great source of wisdom, helping to guide me through the difficulties of being a graduate student in a foreign country.

There are so many people that have played a large part in my perseverance through the doctoral program. I would like to thank Professor Ricardo Eusebi, Professor Stephen Fulling, Professor Simone Gennai, Professor John Cumalat, Professor Will Johns, and Shuichi Kunori for the many great physics talks and advice.

TABLE OF CONTENTS

Page

Page

Page

LIST OF TABLES

TABLE Page

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

FIGURE Page

1. INTRODUCTION

For the last few decades, the Standard Model (SM) of particle physics has served physicists well as a means of understanding the world around us. However, as the scope of experimental observation and data expand to otherwise previously untouched realms of physics, we have begun to challenge our current understanding of the fundamental laws of physics. As physicists, any potential sign of new physics is merely an opportunity to seek and to dream beyond the established models and theories so that we may begin to push knowledge forward.

Physicists and philosophers have been asking the same fundamental questions for hundreds of years:

- What are the parts that make up the world around us and how do those parts interact?
- What is the single event that constitutes the beginning of time and how did the universe evolve from such an event to what we observe today?

Although we may not answer those questions in our lifetime, year by year we are able to piece together some part of the puzzle that is our universe. The evidence that exists to suggest our current understanding of particle physics is incomplete is undeniable. For example, astronomical measurements from the Wilkinson Microwave Anisotropy Probe (WMAP) show that only 4 % of the content of the energy density of the universe can be explained by the SM of particle physics [1].

What is the origin of mass? Why are some particles massless? What is Dark Energy? These are all fundamental questions for which there are no universally

This dissertation follows the style of Journal of Instrumentation.

accepted answers. The Large Hadron Collider (LHC), located at the European Center for Nuclear Research (CERN), is a gigantic particle accelerator built by scientists to answer these unresolved questions. It is the centerpiece of the world's largest experiments and is designed to recreate some of the conditions of the hypothesized Big Bang and probe new physics.

Because this potentially new realm of physics is previously uncharted territory, there is uncertainty as to what this new physics might look like. However, we may use our current understanding of the SM and the many physically well motivated "beyond the Standard Model" (BSM) theories as a basis for establishing new methods to search for these potential physics signatures. The SM has successfully described many experimental observations and even successfully predicted the existence of various particles and their properties. One of the elementary SM particles, the *tauon* or *tau* (τ), was discovered through a series of experiments conducted between 1974 and 1977 at the Stanford Linear Accelerator Center (SLAC) [2]. Taus play an important role in the description and phenomenology of many BSM theories. In this analysis, data from proton-proton (pp) collisions at the LHC is used to perform a general search for new hypothetical massive particles decaying to highly energetic τ lepton pairs.

The thesis is organized as follows: a detailed description of the SM and possible extensions of the SM are described in Chapter 2. The experimental apparatus, the LHC and the Compact Muon Solenoid (CMS) detector, is described in Chapters 5 and 6. In Chapter 7, τ leptons and their importance for physics searches beyond the SM are described. The general search strategy is discussed in Chapter 10, while particle identification and event reconstruction is described in detail in Chapter 13. Finally, the details and results of the analysis are provided in Chapters 14 – 22.

2. PARTICLES AND INTERACTIONS

2.1 The Standard Model

In high energy physics, the best experimentally consistent knowledge of how particles and matter interact is encompassed in the Standard Model (SM). It successfully describes an extraordinarily large amount of experimental data from various experiments. However, there are a growing number of observations that suggest the SM is incomplete. The underpinning of the SM is quantum field theory, a powerful theory that describes the fundamental principles of quantum mechanics and special relativity. The SM successfully incorporates and describes the electromagnetic interactions - which accounts for atomic, molecular, optical, and condensed matter physics - and the strong (also known as quantum chromo dynamics or QCD) and weak interactions responsible for nuclear processes. The gravitational force is the only known fundamental interaction that has not been successfully incorporated into the SM. However, the interaction strength of the gravitational force is small compared to the other three fundamental forces and plays no role in collider experiments.

According to the SM, matter is composed of 12 *fermions* (spin 1/2 particles), while the fundamental forces of nature between fundamental particles and matter (electromagnetic, weak, and strong) are mediated by *bosons* (integer spin particles). SM fermions are divided into two subgroups: *leptons* and *quarks*. Leptons are involved primarily in electroweak interactions, and thus are characterized by an electric charge which couples to the $\gamma/W^{\pm}/Z^{0}$ bosons that mediate the interaction. Leptons are divided into first, second, and third generation fermions. The first generation leptons consist of the electron and corresponding electron neutrino, e and ν_e. The second generation leptons consist of the muon, μ, and the corresponding muon neutrino, ν_μ. Similarly, the tau and the tau neutrino, τ and ν_τ, are the third generation leptons. The most significant difference between generations of leptons are the masses

of the particles. For example, the mass of the electron is 0.511 MeV/c^2 while the mass of the tau is 1777 MeV/c^2, a difference of almost 10^4. Quarks possess electric charge and interact via electroweak interactions; however, they are also characterized by color charge which couples to the gluons that mediate interactions involving the strong force. Like the leptons, quarks are divided into three generations or groups. The up and down quarks, u and d, are the first generation quarks and combine to form the protons and neutrons within atoms. The charm and strange quarks, c and s, are the second generation quarks, while the top and bottom quarks, t and b, are of the third generation.

The forces that govern the interactions between elementary particles and matter are mediated by gauge bosons. The gauge groups of the SM, $SU(3)_C$x$SU(2)_L$x$U(1)_Y$, govern the symmetries and the underlying physics behind the bosons that mediate the forces. For example, the gauge group $SU(3)_C$ contains the underlying symmetries and governs the way the strong force is felt by particles with color charge, such as quarks. The gauge groups of $SU(2)_L$ and $U(1)_Y$ together determine the symmetries and underlying physics involved with the electroweak force. The $W^{\pm}/Z^0$ bosons mediate the weak interactions between particles of different flavors. The weak interactions involving $W^{\pm}$ act on left-handed particles and right-handed anti-particles only, while the Z^0 boson interacts with both left-handed particles and right-handed anti-particles. More specifically, the $SU(2)_L$ gauge group governs the physics of left-handed helicity states of fermion pairs such as (u_L, d_L) and $(e_L, \nu_{e,L})$, while the $U(1)_Y$ gauge group governs the force felt by particles with hypercharge.

The SM also predicts a yet–to–be–discovered boson called the *Higgs* boson. The Higgs boson is responsible for electroweak symmetry breaking via the Higgs mechanism (section 2.3). For now, it is sufficient to say that the mechanism of electroweak symmetry breaking causes the gauge groups $SU(2)_L$x$U(1)_Y$ of the electroweak force

to "break" into the electromagnetic and weak forces. The resulting group, $U(1)_{EM}$ governs the forces felt by the electromagnetic force. Photons, γ, mediate the electromagnetic force between electrically charged particles whereas the other gauge bosons of $SU(2)_L \text{x} U(1)_Y$ form the gauge bosons of $W^{\pm}/Z^0$ that govern the weak interaction.

Matter is merely a composition of the fundamental particles bound together in some physical state only because bosons are able to exchange information about the way they should interact. For example, matter containing 2 or 3 quarks held together by the strong force (gluons) give rise to a new set of particles called Hadrons. Protons are just one example of a Hadron composed of 3 quarks (Bayons). Because the strong interaction does not allow quarks to exist in a "free" state (asymptotic freedom), quarks always exist in bound states. In the same way opposite electric charges can attract to form bounded states, the color charge of quarks is stabilized by binding two or three quarks together resulting in objects/particles known as Mesons (quark and anti-quark combinations) and Baryons (three quarks). Mesons and Baryons are collectively known as Hadrons.

The electric charge of quarks is measured with respect to the electric charge of the electron, with values $(\frac{2}{3}, \frac{-1}{3})|e|$ for each doublet, whereas leptons have electric charge $(-1, 0)|e|$. Similarly, there are three possible values of color charge: "red", "blue", and "green." Additionally, each particle in the SM has a partner anti-particle with the same mass but opposite charge (electric or color). Table 2.1 summarizes the fundamental properties of particles in the SM. The Higgs boson (to be discussed in section 2.3) is yet to be discovered and thus is not listed in Table 2.1.

Table 2.1
SM Particles and their Properties.

Particle	Charge	Spin	Mass
$e^{\pm}$	± 1	$\frac{1}{2}$	$0.51 \pm 0.13\text{x}10^{-9}$ MeV
$\mu^{\pm}$	± 1	$\frac{1}{2}$	105.658 MeV
$\tau^{\pm}$	± 1	$\frac{1}{2}$	1776.84 ± 0.17 MeV
ν_e	0	$\frac{1}{2}$	$<<$ 2 eV
ν_μ	0	$\frac{1}{2}$	$<<$ 0.19 MeV
ν_τ	0	$\frac{1}{2}$	$<<$ 18.2 MeV
u	$\frac{2}{3}$	$\frac{1}{2}$	1.5 to 3.3 MeV
d	$\frac{-1}{3}$	$\frac{1}{2}$	3.5 to 6.0 MeV
c	$\frac{2}{3}$	$\frac{1}{2}$	$1.27^{+0.07}_{-0.11}$ MeV
s	$\frac{-1}{3}$	$\frac{1}{2}$	104^{+26}_{-34} MeV
t	$\frac{2}{3}$	$\frac{1}{2}$	172.2 ± 2.1 GeV
b	$\frac{-1}{3}$	$\frac{1}{2}$	$4.2^{+0.17}_{-0.07}$ GeV
γ	$< 5\text{x}10^{-30}$	1	$1\text{x}10^{-18}$ eV
$W^{\pm}$	± 1	1	80.398 ± 0.025 GeV
Z^0	0	1	91.1876 ± 0.0021 GeV
g	0	1	0

2.2 Symmetry Breaking

One of the beautiful aspects of the SM is the intrinsic symmetries of the model through the unitary groups: U(1), SU(2), SU(3). Symmetry under U(1) means the fields are invariant under phase transformations $\phi \rightarrow \phi' = e^{i\theta}\phi$, so that the physical properties arising from the SM are not affected. Similarly, symmetry under the special unitary groups SU(2) and SU(3) means the physical state is unaffected by rotations in the two- and three dimensional complex space respectively. The $SU(3)_C\text{x}SU(2)_L\text{x}U(1)_Y$ structure of the SM is depicted in Figure 2.1. The quarks are arranged in triplets with respect to the color gauge group $SU(3)_C$, with indices as red (r), green (g), and blue (b).

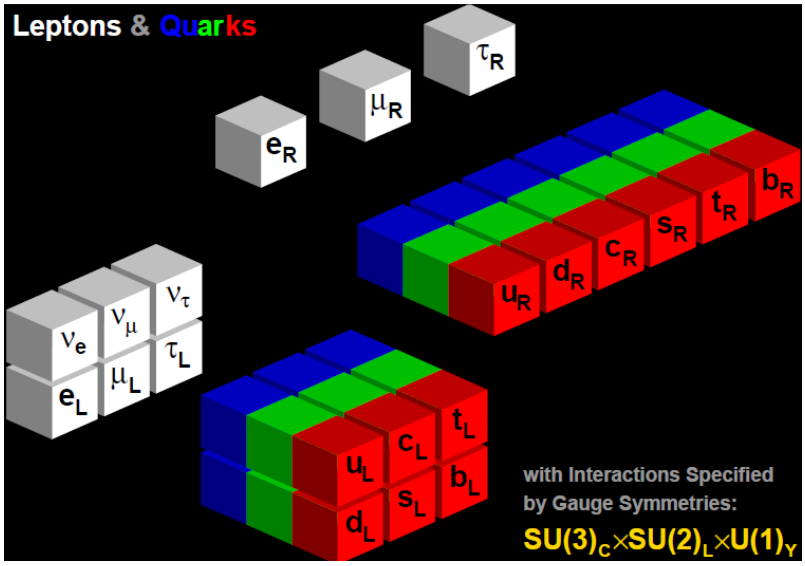

Fig. 2.1. $SU(3)_C$x$SU(2)_L$x$U(1)_Y$ structure of the Standard Model [3].

$$q = \begin{pmatrix} q_r \\ q_g \\ q_b \end{pmatrix}$$

The left- and right-handed fermions have different transformation properties under the weak isospin group $SU(2)_L$. The left-handed fermions are arranged in doublets, and the right-handed fermions are arranged in singlets:

$$Leptons : \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \quad e_R \quad \mu_R \quad \tau_R$$

$$Quarks : \begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L \quad u_R \quad d_R \quad c_R \quad s_R \quad t_R \quad b_R$$

The electric charge of fermions is related to the hypercharge Y of $U(1)_Y$ and weak isospin T of $SU(2)_L$ by $Q = T_L^3 + \frac{Y}{2}$. The quantum numbers for the first generation fermions are listed in Table 2.2, and the couplings to fermions in the SM are listed in Table 2.3 [3].

The SM postulates that a physical state must be invariant under a space-time dependent phase transformation. With this assumption alone, the mediators of the fundamental interactions must be massless. However, experimental evidence suggests otherwise. Although this is true for the photon, the Z and W bosons are known to have masses of 91.19±0.0021 GeV/c^2 and 80.40±0.025 GeV/c^2 respectively [4]. The measurement of the massive $W^{\pm}/Z^0$ masses changed the scope of particle physics

Table 2.2
First Generation Fermion Quantum Numbers [3].

Particle	Q	T_L^3	Y	C
ν_e	0	1/2	-1	0
e_L	-1	-1/2	-1	0
e_R	-1	0	-2	0
u_L	2/3	1/2	1/3	r,g,b
d_L	-1/3	-1/2	1/3	r,g,b
u_R	2/3	0	4/3	r,g,b
d_R	-1/3	0	-2/3	r,g,b

Table 2.3
Couplings to Fermions in the Standard Model.

	Left Coupling	Right Coupling
Strong: $g \to q\bar{q}$	$\frac{g_3}{2}\lambda^a$	$\frac{g_3}{2}\lambda^a$
Electromagnetic: $\gamma \to f\bar{f}$	eQ_f	eQ_f
Weak: $Z^0 \to f\bar{f}$	$\frac{g_2}{cos\theta_W}(T_f^3 - sin^2\theta_W Q_f)$	$-\frac{g_2}{cos\theta_W}sin^2\theta_W Q_f$
Weak: $W^\pm \to l\nu_l$	$\frac{g_2}{\sqrt{2}}$	0
Weak: $W^\pm \to qq'$	$V_{qq'}\frac{g_2}{\sqrt{2}}$	0

and eventually led to the development of the Higgs mechanism which we believe to be responsible for the coupling of interactions to mass. Spontaneous symmetry breaking is the mathematical outcome when a theory has degenerate ground state solutions. Choosing one of the degenerate solutions "breaks" the symmetry of the system. The SM accounts for the discrepancy between the observed massive gauge bosons and the previously hypothesized massless states by stating the "broken symmetry" occurs in the vacuum state where a still-unobserved field, the Higgs field, allows for these observed mass eigenstates.

In general, adding terms of the form $m^2VV^\dagger$, where V denotes the Z and W gauge fields, to the Lagrangian gives rise to massive gauge bosons, but does not preserve

the gauge symmetry. One can understand the problem of broken symmetries by considering a $U(1)$ gauge theory with a single field, the photon [5]. The Lagrangian can be written as:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \tag{2.1}$$

where

$$F_{\mu\nu} = \partial_\nu A_\mu - \partial_\mu A_\nu \tag{2.2}$$

such that A is the vector potential. Local $U(1)$ gauge invariance means that the Lagrangian is invariant under the transformation: $A_\mu(x) - \partial_\mu \eta(x)$ for any η and x. If we add a term that were to represent the mass of the photon, the Lagrangian can become:

$$L = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} m^2 A_\mu A^\mu \tag{2.3}$$

However, this new Lagrangian now violates the rule of local gauge invariance. As stated above, our physical model or mathematics that describes the fundamental interactions must satisfy this rule of local gauge invariance; otherwise, we could easily transform our physical state to a different space-time coordinate and arrive at physical paradoxes. Therefore, this is one example of how local gauge invariance can dictate either massive or massless gauge bosons. In the case presented above, the $U(1)$ gauge invariance dictates that the photon be massless.

One can consider an extension of the above case in which a single complex scalar field is added to represent the coupling of charge to the photon. In this case, the Lagrangian becomes

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + |\partial_\mu\phi - ieA_\mu\phi|^2 - \mu^2|\phi|^2 - \lambda(|\phi|^2)^2 \tag{2.4}$$

This Lagrangian is invariant under $U(1)$ rotations, $\phi \rightarrow e^{i\theta}\phi$, and also under the gauge transformations $A_\mu(x) - \partial_\mu\eta(x)$ and $\phi(x)e^{-ie\eta(x)}$. The Lagrangian describes quantum electrodynamics with a massless photon and a charge scalar field ϕ with mass μ [5].

2.3 Higgs Mechanism

The Higgs particle is a hypothetical massive scalar elementary particle with no intrinsic spin that plays the unique role of explaining why the photon has no mass, while the W and Z bosons are massive. The Higgs field is introduced within the SM to account for the broken $SU(2)_L$x$U(1)_Y$ symmetry. The Lagrangian density for this field is given by:

$$l = (\partial_\mu\phi)^\dagger(\partial^\mu\phi) - V(\phi) \tag{2.5}$$

where the potential $V(\phi)$ is given by:

$$V(\phi) = \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 \tag{2.6}$$

and $\lambda > 0$ is required to maintain bounded potentials as $\phi \to \infty$. The SM Higgs field is expressed as a doublet in terms of two complex fields:

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

where:

$$\phi^+ = \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \phi^0 = \frac{\phi_3 + i\phi_4}{\sqrt{2}} \tag{2.7}$$

In order to determine the vacuum state, the Lagrangian can be minimized under the assumption of an Abelian Higgs field (representation in which the field is described in terms of one scalar and one imaginary component). Minimization in this way results in:

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{\nu^2}{2} \tag{2.8}$$

For $\mu^2 < 0$, the potential has two stable ground states, $+\nu$ and $-\nu$. This is depicted in Figure 2.2 and is often referred to as the "mexican hat" potential. If one chooses $+\nu$ or $-\nu$ as the ground state and defines the direction of the field, one obtains:

$$\phi_3 = \nu, \phi_1 = \phi_2 = \phi_4 = 0 \to \phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu \end{pmatrix} \tag{2.9}$$

By expanding around one of the ground states, the excited states of the field can be determined. If the perturbation is $H(x)$, then

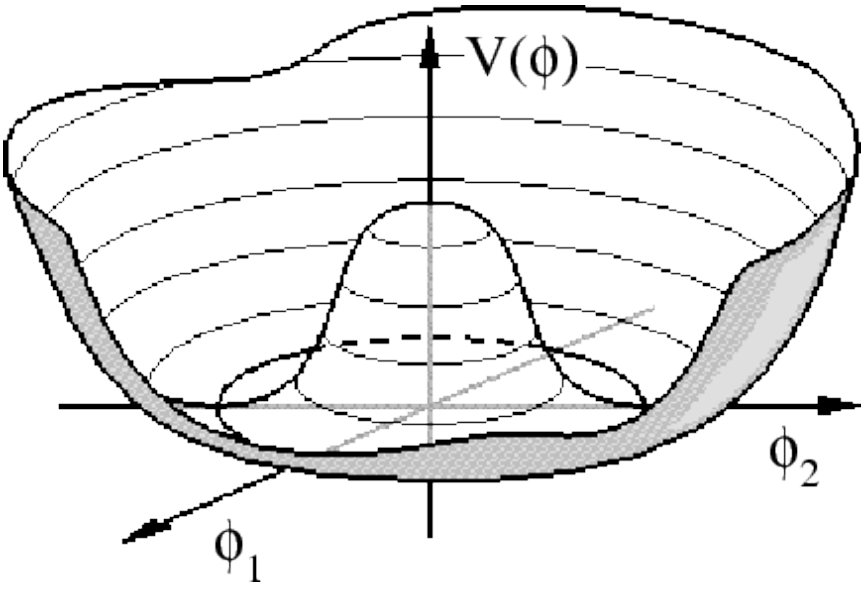

Fig. 2.2. Higgs Potential.

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ \nu + H(x) \end{pmatrix} \tag{2.10}$$

These excited states are the particles within the SM. Choosing the direction of the field breaks the symmetry; however, introducing a covariant derivative allows for the preservation of the $SU(2)_L \mathrm{x} U(1)_Y$ symmetry, while also accomodating massive gauge bosons [5]. The accurate mass prediction of the $W^{\pm}$ and Z^0 bosons within this model is one of the most significant accomplishments in physics. It is believed that a large amount of energy and luminosity is required to observe a Higgs boson (with a certain degree of confidence) at high energy colliders. For this reason, it is hypothesized that the LHC will confirm the existence of this particle.

3. MOTIVATION FOR Z' BOSONS

Much like the Higgs mechanism was introduced to account for the $SU(2)_L$x$U(1)_Y$ symmetry breaking, one can construct theoretical models by incorporating additional gauge fields that can lead to heavy gauge bosons yet to be discovered. Although there are several ways in which new heavy gauge bosons appear, the most natural possibility is one in which these heavy gauge bosons are simply the gauge field of a new local broken symmetry. Examples include models with a new U(1) gauge symmetry, little Higgs models, and E_6 Grand Unified Theories (GUT) [6].

Z' bosons refer to hypothetical new gauge bosons that arise from extensions of the electroweak symmetry of the Standard Model. In models with a new U(1) gauge symmetry, the Z' is the gauge boson of the broken symmetry. In little Higgs models there is typically an enlarged gauge sector that is broken down to the SM gauge symmetry around the TeV scale that results in one or more Z' bosons. In Kaluza-Klein models, the Z' bosons are excited states of a neutral bulk gauge symmetry.

In order to show how these new mass eigenstates can present themselves, first consider the Lagrangian for the neutral current interactions of fermions, J, in the SM:

$$L = gJ_3^\mu W_{3\mu} + g'J_Y^\mu B_\mu \tag{3.1}$$

The $SU(2)_L$ and $U(1)_Y$ gauge couplings are g and g' respectively, while $W_{3\mu}$ is the $SU(2)_L$ gauge boson and B_μ is the $U(1)_Y$ gauge boson. Considering the symmetry breaking of $SU(2)_L$x$U(1)_Y$, the mass eigenstate neutral gauge bosons are the massless photon field A_μ and the massive Z_μ given by

$$A_\mu = sin\theta_W W_{3\mu} + cos\theta_W B_\mu \tag{3.2}$$

$$Z_\mu = cos\theta_W W_{3\mu} - sin\theta_W B_\mu \tag{3.3}$$

with the weak angle $\theta_W = tan^{-1}(g'/g)$ and new gauge couplings $e = gsin\theta_W$ and $g_1^2 = g^2 + g'^2 = g^2/cos^2\theta_W$. In the extension to $SU(2)_L \mathrm{x} U(1)_Y \mathrm{x} U(1)'^n$ (the simplest scenario in which hypothetical Z' bosons are considered), the Lagrangian becomes

$$L = eJ^\mu_{EM} A_\mu + \sum g_\alpha J^\mu_\alpha Z^0_{\alpha\mu} \tag{3.4}$$

where $\alpha = 1 \rightarrow n+1$ and g_1, $Z^0_{1\mu}$, and J^μ_1 are the gauge coupling, boson, and current of the SM. Additionally, g_α and $Z^0_{\alpha\mu}$ are the gauge couplings and bosons of the additional $U(1)'$ extension. If one assumes the electrically neutral scalar fields attain vacuum expectations values such that A_μ remains massless, then the $Z^0_{\alpha\mu}$ fields acquire mass terms $\frac{1}{2} M^2_{\alpha\beta} Z^0_{\alpha\mu} Z^{0\mu}_\beta$ such that

$$M^2_{\alpha\beta} = 2 g_\alpha g_\beta \sum Q_{\alpha i} Q_{\beta i} |\phi|^2 \tag{3.5}$$

where $M_{11} = M^2_{Z^0}$ would be the SM Z mass [6]. Diagonalizing the mass matrix, one obtains $n+1$ massive eigenstates $Z_{\alpha\mu}$ with mass M_α. The simplest scenario in which new massive eigenstates are generated has been considered; however, the underlying mathematics and physical motivation behind other extensions of the SM work in a similar fashion.

Historically, physicists have adopted the SM as the preferred model of particle physics because it is the simplest model that accounts for experimental observations. However, as the scope of high energy experiments expands to increasing energy regimes, it is likely nature will reveal itself in other forms. In fact, there is already evidence to suggest the SM is not the correct model (e.g. possible Dark Matter signatures from astronomical data [1]). In addition to searches for the Higgs bosons, hypothetical Z' bosons from arising from new gauge symmetries is another example of new massive resonances that can be searched for at the LHC.

4. HIGH MASS RESONANCES DECAYING TO TAU PAIRS

One of the main SM τ pair production mechanisms at the LHC is from Drell-Yan processes $pp \to \gamma^*/Z \to \tau\tau$. The leading order Feynman diagram for the production of $Z \to \tau\tau$ is shown in Figure 4.1. The production mechanism occurs mainly through the annihilation of a quark and an anti-quark, resulting in an intermediate Z boson that decays to a pair of outgoing τ leptons.

Higgs bosons provide another means by which one can observe τ pair production at the LHC. For example, there is a large probability that the pseudoscalar Higgs boson A predicted in minimal supersymmetric extensions of the SM (MSSM) can only be discovered/observed in cases where it decays to a pair of τ leptons (more details in section 7.3). This can be shown be considering the Born approximation of the Higgs boson decay width [6]. The width into charged lepton pairs is

$$\Gamma(A \to l^+ l^-) = \frac{G_F M_l^2}{4\sqrt{2}\pi} M_A \beta_l^3 \tag{4.1}$$

where $\beta_l = \sqrt{1 - 4M_l^2/M_A^2}$ is the velocity of the final state leptons, while the decay width into quarks

$$\Gamma(A \to q\bar{q}) = \frac{3G_F M_q^2}{4\sqrt{2}\pi} M_A \beta_l^3 (1 + \frac{4}{3}\frac{\alpha_s}{\pi}\delta_A^{QCD}) \tag{4.2}$$

is enhanced by a color factor of 3 and also receives significant QCD corrections. Because of the dependence on M_l^2, if one considers leptonic decays, the Higgs boson decays predominantly to pairs of τ's. Due to the presence of the QCD correction

factor for Higgs decays to quarks, in some cases (depending on the Higgs mass) the Higgs boson can decay predominantly to quarks. However, the inclusive search for $pp \to A \to b\bar{b}$ is difficult due to the large amount of backgrounds from $b\bar{b}$ production. This means the $\tau\tau$ mode can be the most promising means for discovery of the Higgs boson. The leading order diagram for possible production of A's is shown in Figure 4.2.

As described in the previous sections, many extensions of the SM provide the possibility for new heavy gauge bosons, Z'. One can consider the simplest $U(1)$ extension in which the Z' has the same interactions as the SM Z boson, the sequential Z', and the only unknown parameter is the mass of this new gauge boson. The tree level Feynman diagram for the production of this sequential Z' is shown in Figure 4.3. One can use the same couplings to τ's as the SM Z boson in order to calculate the cross-section. The predicted cross-section as a function of the sequential Z' mass is shown in Figure 4.4 and Table 4.1. The Higgs branching ratios as a function of Higgs mass are shown in Figure 4.5.

In this analysis, we perform a direct search for new hypothetical massive (> 100 GeV) particles decaying to a pair of τ leptons $X \to \tau\tau$. We take the sequential Z' as our benchmark model to construct a robust, optimal analysis and calculate the signal acceptance for a generic kind of new massive scalar boson.

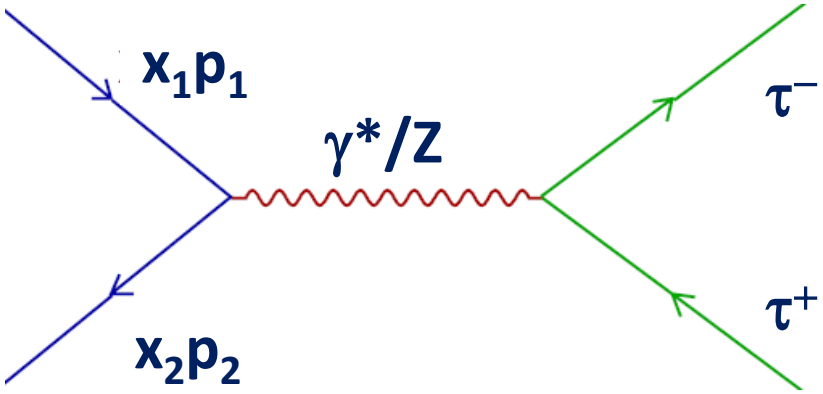

Fig. 4.1. Standard Model Drell-Yan production.

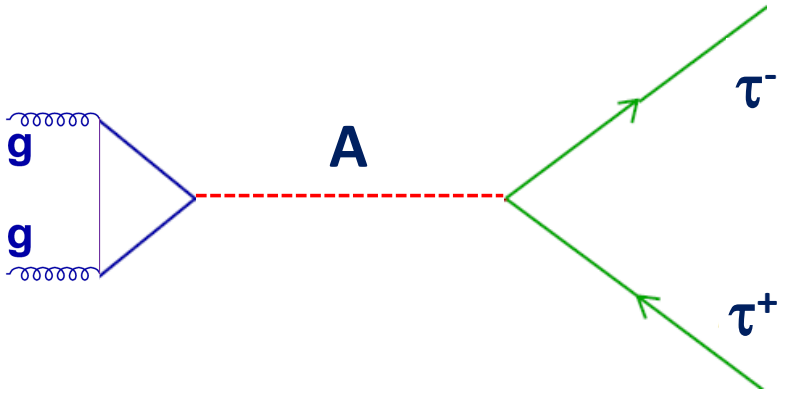

Fig. 4.2. MSSM Pseudoscalar Higgs production.

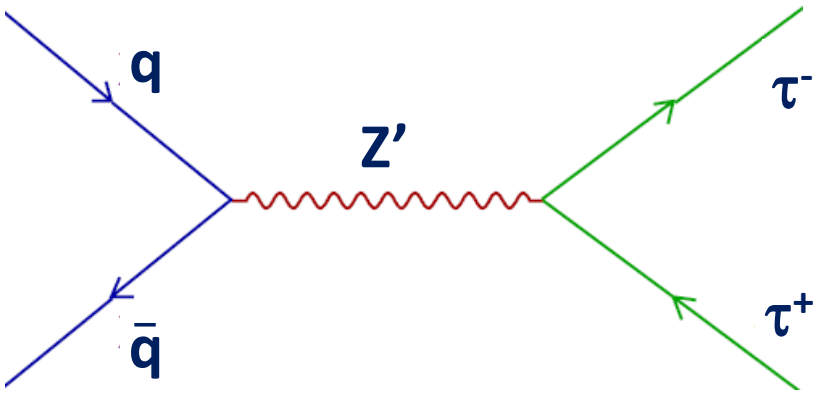

Fig. 4.3. Sequential Z' production.

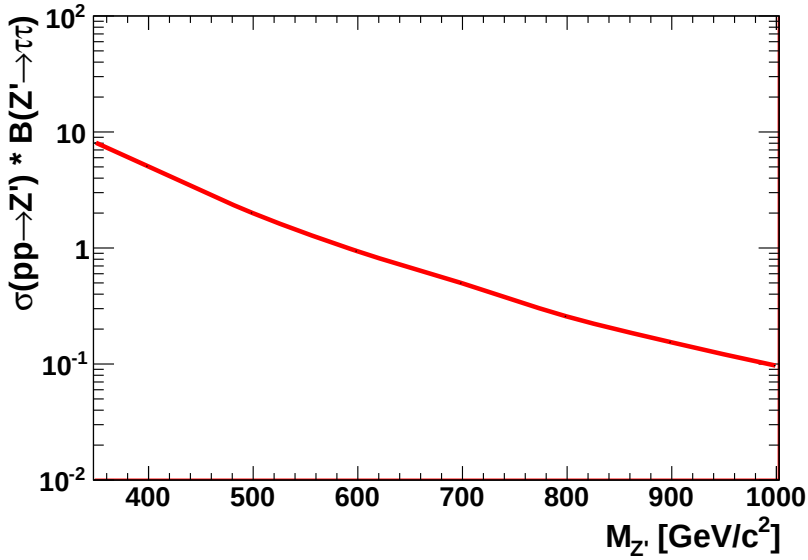

Fig. 4.4. Theoretical signal sequential Z' cross-section times the branching ratio to τ's as a function of mass.

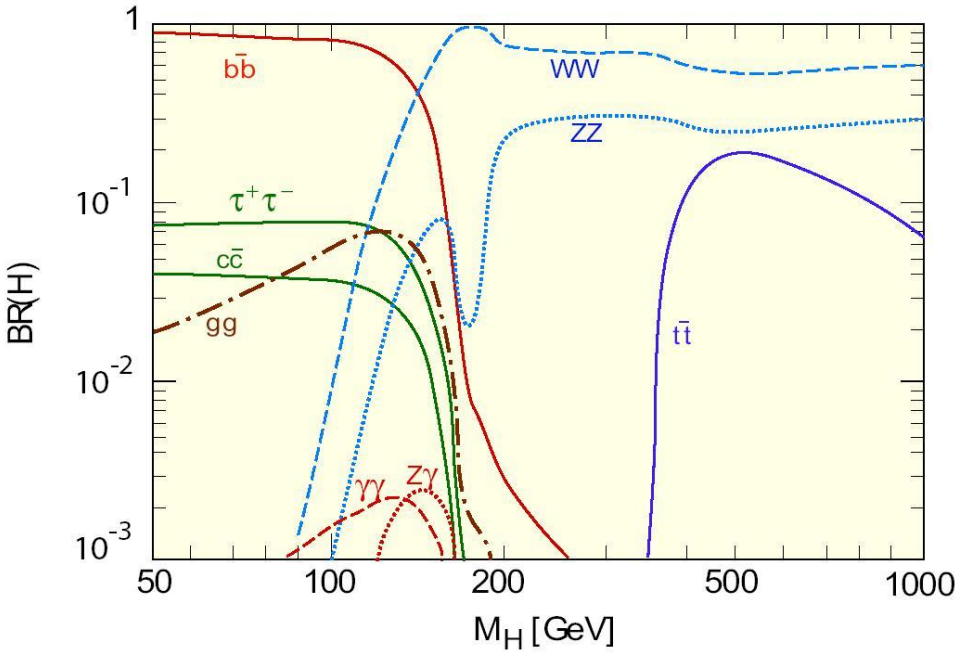

Fig. 4.5. Higgs Branching Ratios as a function of Higgs Mass.

Table 4.1
Sequential Z' Cross-section times the Braching Ratio to Taus For Various Masses.

Mass (GeV/c^2)	$\sigma(pp \to Z')$ x BR($Z' \to \tau\tau$) (pb)
350	8.061
400	4.980
500	1.974
600	0.930
700	0.489
800	0.255
900	0.151
1000	0.096

5. THE LARGE HADRON COLLIDER

The Large Hadron Collider (LHC) is a two-ring superconducting proton-proton particle accelerator built at the European Center for Nuclear Research (CERN), located on the French-Swiss border near Geneva, Switzerland. The LHC is an underground circular machine, approximately 100 meters underground, spanning 27 km in circumference built to recreate some of the conditions of the hypothesized Big Bang and probe new physics. The motivation behind the design of the LHC is to determine the nature of electroweak symmetry breaking for which the Higgs mechanism is presumed to be responsible. It is also believed that the LHC could shed light as to the mathematical consistency of the SM at energy scales of 1 TeV.

The CERN accelerator complex has two transfer tunnels, each approximately two and a half kilometers in length, linking the LHC to the CERN accelerator complex that acts as an injector. Protons are accelerated to energies of 50 MeV by the linear accelerator called LINAC2. The protons are then injected in to the Proton Synchrotron Booster (PSB) which utilizes radio frequency (RF) cavities to accelerate the protons to an energy of 1.4 GeV. The protons then go through the Proton Synchrotron (PS) and are accelerated to an energy of 25 GeV. Finally, the Super Proton Synchrotron (SPS) accelerates the protons to energies of 450 GeV before being sent to the LHC.

As the protons travel through the LHC complex, they are accelerated via radio frequency cavities to velocities close to the speed of light. Dipole magnets are used to guide the beam of protons around the ring. The dipole magnets are built from coils of special electric cable that operates in a superconducting state, efficiently conducting electricity without resistance or loss of energy. To achieve this, the magnets must be cooled to extremely low temperatures, approximately -270 degrees C, lower than the

measured temperatures of outer space. Because the probability that two subatomic particles will interact as they travel toward each other is very small, protons are placed in "bunches" and these bunches are smashed into each other so as to increase the probability that two of the protons within the oppositely moving bunches will interact. Furthermore, the beam of bunches are "squeezed" so that the protons within the bunches are situated within some smaller cross-section to additionally increase the probability of an interaction to occur. The "squeezing" of the beams is accomplished with quadrapole magnets situated around the LHC complex, each 5-7 m long.

Luminosity is a measure of the probability that a proton will interact/collide with another proton. High luminosities can be achieved in several ways: (1) increasing the number of protons that can be placed within each bunch; (2) ensuring that the beams are properly "squeezed" into collimated beams of small cross-sectional area; and (3) increasing the bunch crossing rate or frequency. More specifically, the instantaneous luminosity is the number of interactions per unit area and unit time. The units of instantaneous luminosity are $b^{-1}\ s^{-1}$ or $cm^{-2}s^{-1}$. A barn, b, is a unit of area satisfying the following conversion: $10\ nb = 10^{34}\ cm^2$. The luminosity integrated over a specific range of time is referred to as the integrated luminosity and is given in units of b^{-1}. Put simply, the integrated luminosity is a measure of the amount of collected data. The LHC luminosity depends on the proton beam parameters and can expressed as:

$$L = \frac{N_b^2 n_b f \gamma_r}{4\pi\epsilon_n \beta^*} F \tag{5.1}$$

where N_b is the number of particles per bunch, n_b the number of bunches per beam, f the frequency of revolution, γ_r the relativistic factor for particles traveling at velocities near the speed of light, ϵ_n the transverse beam emittance, β^* the beta function

at the collision point, and F the geometric factor due to the crossing angle at the interaction point [7]. The design luminosity of the LHC is $10^{34}\ cm^{-2}s^{-1}$ with approximately 2,800 bunches per proton beam, and a bunch spacing of 25 ns leading to approximately 10^9 inelastic events/s.

It is believed the reason we have not discovered the Higgs boson and/or have not observed signs of new physics beyond the SM is because previous experiments have not attained the energies necessary to produce such massive particles. Additionally, even if one assumes that the provided energies are sufficient, the rate at which these new physics events can be produced is significantly smaller than the already well established SM processes. The total number of events produced in a scattering process is directly proportional to the integrated luminosity, L, and the scattering cross-section σ of the process:

$$N = L\sigma \tag{5.2}$$

The scattering cross-section is a measure of the effective area seen by impinging particles. Said differently, a scattering cross-section is a hypothetical area that quantifies the likelihood for impinging particles to interact. To get a rough idea of the challenge of designing an analysis that is sensitive to new physics signatures, one can consider the cross-sections (at center of mass energy of 7 TeV) for the following examples:

$$\sigma(pp \to Z \to \tau\tau) \sim 1650\ pb \tag{5.3}$$

$$\sigma(pp \to t\bar{t}) \sim 150\ pb \tag{5.4}$$

$$\sigma(pp \to Z'(500) \to \tau\tau) \sim 2\ pb \tag{5.5}$$

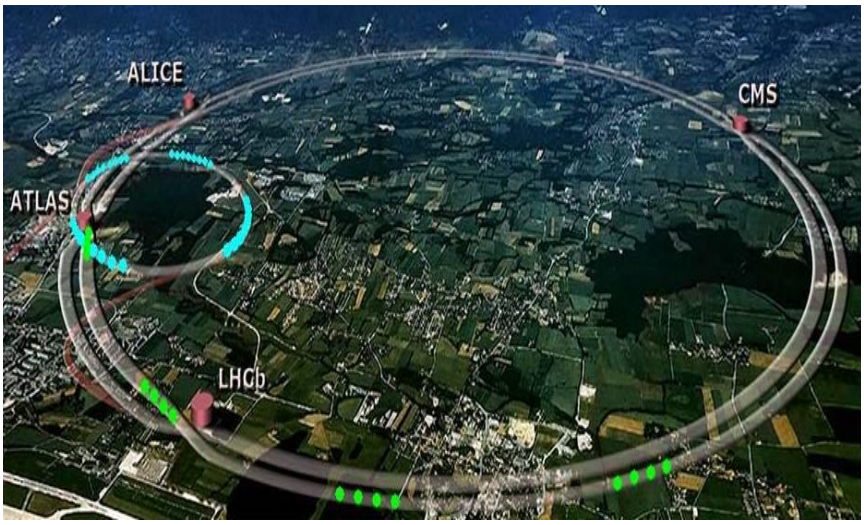

Fig. 5.1. Aerial View of the LHC [8].

From these examples, it is clear that attempting to distinguish e.g. a sequential Z' signature from the copious amounts of SM processes requires a rejection factor of at least 10^{-3}.

The LHC began producing 7 TeV pp collisions in May of 2010 and delivered approximately 45 pb^{-1} of pp collision data. In this thesis, we make use of the 36.15 pb^{-1} of Run 2010A and Run 2010B data collected by the Compact Muon Solenoid (CMS) detector between May and December of 2010. An aerial view of the LHC is displayed in Figure 5.1. The CERN accelerator complex is depicted in Figure 5.2.

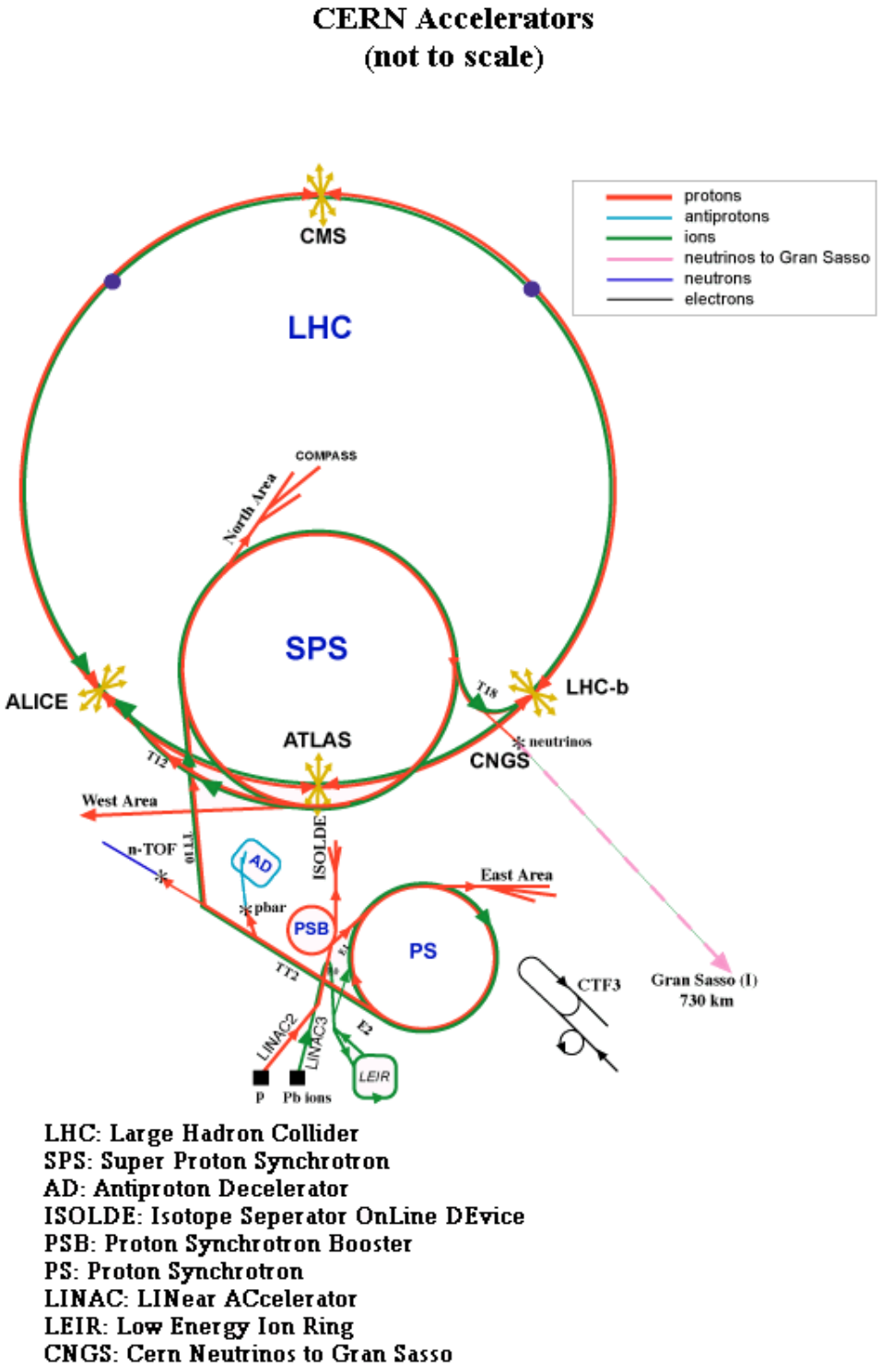

Fig. 5.2. CERN Accelerator Complex [9].

6. THE CMS DETECTOR

The LHC has six multi-purpose detectors designed to probe a wide variety of physics processes. The Compact Muon Solenoid (CMS) and ATLAS [7] detectors are the two largest experiments and are expected to search for very similar physics signatures including the Higgs mechanism described in sections 2.3 and 2.2, the possible nature of dark matter, and insight into Grand Unified Theories such as minimal Supergravity (mSUGRA). In this section, focus is placed on the properties of the CMS detector and how it relates to the detection of particles and search for new physics.

The goals of the LHC physics program dictates the requirements for the design of the CMS detector. Good muon identification and resolution is necessary in order to achieve dimuon mass resolutions of $\sim 1\%$ at 100 GeV, where the Higgs boson is expected to create dimuon signatures. Similary, good electromagnetic energy resolution is required to produce diphoton and dielectron mass resolutions on the order of $\sim 1\%$ at 100 GeV. Finally, calorimeters with large hermetic geometric coverage are necessary to achieve good missing transverse energy resolution (section 13.6) that will be vital for the discovery of new physics signatures such as Supersymmetry (SUSY).

At design LHC luminosities, the bunch crossing frequency is 40 MHz, which means billions of pp interactions per crossing are produced. However, we do not have the capabilities to store information from every single interaction. Therefore, hardware has been designed to select interesting events and achieve a manageable rate. At the first trigger level (L1), very basic selections are employed to select events that contain certain particle signatures with sufficient energies and reduce the event rate to 100 kHz. At the second level, High Level Trigger (HLT), more complex al-

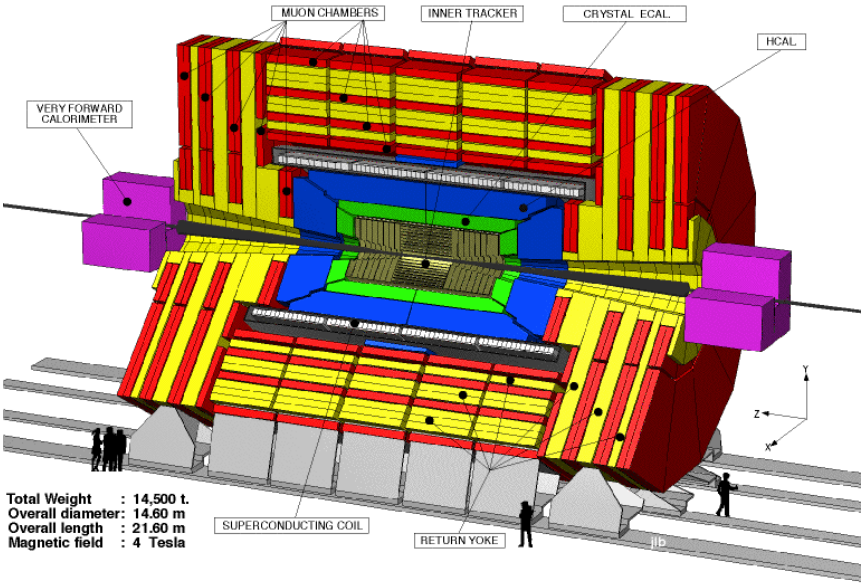

Fig. 6.1. CMS Detector.

gorithms are used to provide more accurate measurements so they may be used to further reduce the event rate to 100 Hz. Finally, events passing the HLT are stored and reconstructed using the best available algorithms to be used at the analysis level.

Particles produced in the *pp* collisions first interact with a silicon tracker subsystem that charts the position as they move through the detector. Outside the tracker sub-detector sit calorimeters designed to measure the energies of electrons, photons, and hadrons. The tracker and calorimeter sub-detectors sit inside the solenoid coil geometry. Finally, the muon system sits outside the inner magnet geometry and is designed to measure the momentum and position of muons. A sketch of the CMS detector is shown in Figure 6.1.

6.1 CMS Coordinate System

The CMS experiment uses a right-handed coordinate system, with the origin at the nominal interaction point, the x-axis pointing to the center of the LHC ring, the y-axis pointing up (perpendicular to the LHC plane), and the z-axis along the anticlockwise-beam direction. CMS uses a spherical coordinate system with the polar angle, θ, measured from the positive z-axis, the azimuthal angle, ϕ, measured in the x-y plane, and the radius r measured from the center of the beamline. The pseudorapidity is given by $\eta = -\ln(\tan\theta/2)$ and is used more often than the polar angle θ. Particles from pp collisions are more likely to "spray" out in the forward regions than in the central regions along $z = 0$. The particle occupancy can be quantified in terms of θ such that

$$d\eta = -\frac{d\theta}{\sin\theta} \tag{6.1}$$

A constant region in η space corresponds to varying regions in θ space that are smaller in the forward regions when compared to the central regions. Therefore, η provides a more uniform description of particle occupancy over the vast geometrical range of the detector.

6.2 Superconducting Magnet

One of the most important features of the CMS detector is the solenoid superconducting magnet which produces a constant magnetic field when electricity flows through it. It is 12.9 m in length and consists of 2,168 turns and an inner bore of 5.9 m which gives rise to a field strength of 4 T, about 100,000 times that of the earth's field strength. Because the typical momenta of particles at the LHC is high, the magnetic field strength is an essential feature required to perform a precise

measurement of charged particle momenta. Due to the number of ampere turns N required to generate a field of 4 T,

$$B \sim \mu \frac{N}{L} I \tag{6.2}$$

the most distinctive feature of the solenoid magnet is a 4-layer winding made from stabilized reinforced NbTi conductor. The important parameters of the CMS magnet are summarized in Table 6.1 [7].

Table 6.1
Parameters of the CMS Magnet [7].

Parameter	Description
Magnetic length	12.5 m
Cold bore diameter	6.3 m
Central magnetic induction	4 T
Ampere turns	41.7 MA-turns
Nominal current	19.14 kA
Inductance	14.2 H
Stored energy	2.6 GJ
Radial thickness of cold mass	312 mm
Weight of cold mass	220t
Maximum induction on conductor	4.6 T
Temperature margin wrt operating temperature	1.8 K
Stored energy/unit cold mass	11/6 kJ/kg
Outer diameter of the iron flats	14 m
Barrel length	13 m
Total mass of iron in return yoke	10,000t

6.3 The Tracker System

One of the most important aspects of searches for new physics consists of identifying particles with a given momentum. The momentum of particles helps build up

a picture of how one believes a collision event originated or of what one believes such an event consists. The main method used to calculate the momentum of a particle is to track its path as it is altered by a magnetic field. According to Newtons's force law, and by using the equation for the Lorentz force on a charged particle in a constant magnetic field, one can determine the momentum of the charged particle:

$$F = ma = m\frac{v^2}{R} = qvBsin\theta \tag{6.3}$$

where R is the radius of curvature, B is the magnitude of the magnetic field, and θ is the angle between the direction of the momentum of particle and the direction of magnetic field. The tracker sub-detector is used to reconstruct the paths of charged particles as they traverse the magnetic field. It is important to note the above equation assumes an unaltered trajectory. Therefore, an accurate measurement of particle momenta requires the tracker material disturb the particle as little as possible.

Many of the most interesting new physics signatures require extremely high luminosities and beam conditions. This can put very stringent demands on the tracking system due to the superposition of many uncorrelated interactions within each bunch crossing. For example, at design luminosities, an average of approximately 20 inelastic collisions can be superimposed on the hard scattering event of interest. Furthermore, approximately 1000 charged particles can emerge from the interaction region every 25 ns. Due to this high multiplicity and density of charged particles, the interaction or event of particular interest may be confused with those from other interactions in the same bunch crossing. This problem can become very severe, but can be tackled by using detectors with high granularity and good time resolution. Distinguishing a hard scattering event from the softer pile-up (PU) interactions requires good separation of interaction positions (vertex). Therefore, good vertex reconstruc-

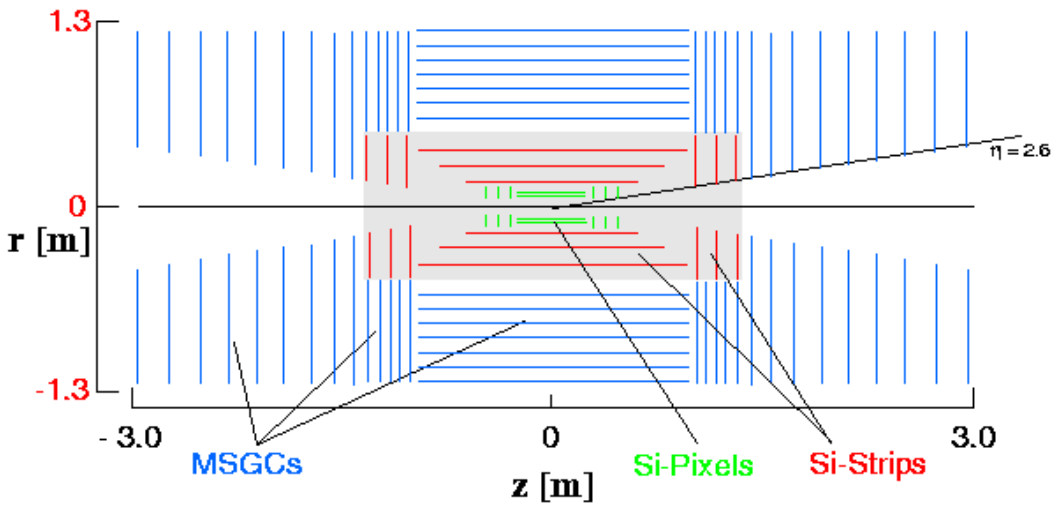

Fig. 8.1. Cross – section of the Tracker.

Fig. 6.2. Cross-sectional view of the CMS Tracker.

tion is an essential part of the tracker design and emphasis must be placed on the ability to recognize complex patterns. When a charged particle interacts with the tracker material, it can bremsstrahlung and cause significant loss in efficiency and reconstruction. This imposes sever constraints on the allowed material budget. Both the efficiency and resolution with which charged particles can be reconstructed with the tracker sub-detector is highly dependent on the material in the tracker.

The tracker system is made entirely of silicon and is divided into two regions: the innermost pixel detector and the micro-strip detectors that surround it. In the central rapidity regions, the tracker detectors are arranged in a barrel geometry, while at higher values of rapidity they are arranged as disks, organized into end-cap like structures. In total there are 13 barrel layers, and the CMS Tracker geometry, shown in Figure 6.2, has been designed so as to provide typically 13 distinct high resolution measurement planes for tracks up to $|\eta|$ of about 2.0, gradually falling to a minimum of 8 planes at $|\eta|$ of 2.5 [7].

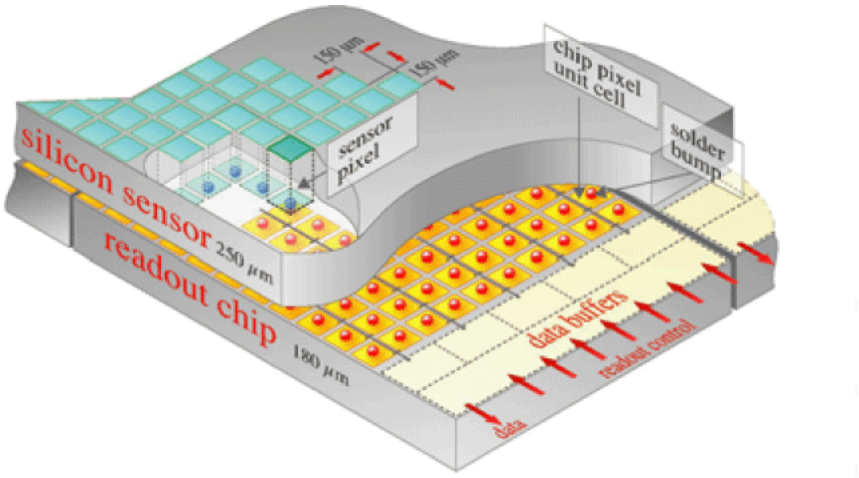

Fig. 6.3. CMS Pixel Detector Layout.

6.3.1 The Pixel Detector

The pixel detector is the closest detector to the beam pipe, which allows it to track the paths of particles with high accuracy. Even more importantly, it allows for the reconstruction of primary and secondary vertices of long lived particles such as b/c quarks or tau leptons. The pixel detector contains about 65 million pixels, each one 100 μm x 150 μm, that produce electron-hole pairs when charged particles pass through them. The charge from the electron-hole pairs is collected as a small electric signal that is amplified by an electronic silicon readout chip (ROC) soldered to each pixel/sensor. The pixel detector consists of three such layers in the barrel region located at mean radii of 4.4, 7.3, and 10.2 cm, and two endcap disks assembled in a turbine-like geometry extending from 6 to 15 cm in radius. The total coverage of the pixel detector extends to $|\eta| < 2.5$. The hit resolutions in the inner pixel detector are approximately 10 μm in the barrel regions, while they can be larger in the endcap pixels (e.g. 50 μm). The CMS pixel detector layout is displayed in Figure 6.3.

6.3.2 The Silicon Strip Detector

After interacting with the silicon pixels, particles produced in collisions interact with layers of silicon strips that extend out to a radius of 130 cm. There are a total of 10 layers of silicon strip detectors, four inner barrel shell layers (TIB) with two inner endcap disks (TID) and six outer barrel layers (TOB). The TIB uses silicon sensors of 320 μm thickness and the TOB uses silicon sensors of 500 μm thickness [7]. The silicon strip detector works in much the same way as the inner pixel detector: as charged particles traverse the material, electrons are knocked off the atoms. These electrons provide a small current due to the applied electric field. This current is amplified and gives rise to what we call "hits", which are merely the amplified electrical signature that is produced when charged particles ionize the atoms within the material. The current is amplified by Analogue Pipeline Voltage (APV) chips which store the signals in memory for a few microseconds before sending them to a laser that converts the signal to infrared pulses. These infrared pulses are then transmitted through fiber optic cables so that they may be analyzed in a radiation free environment [7].

6.4 The Muon System

According to theory, if the Higgs boson exists, it is likely to decay via $H \rightarrow ZZ \rightarrow llll$ where l is a lepton. Therefore, good reconstruction of muons is an essential aspect of physics with the CMS detector. The muon system is composed of three gaseous detectors: drift tube chambers (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC). The muon detectors are interleaved with a 12-sided iron structure that surrounds the magnet with the purpose of containing and guiding the magnetic field. The DT system, spanning the barrel region of the CMS detector, consists of 4 cm wide tubes that contain a wire within the gas volume. When a muon strikes the drift tube, the gas is ionized and the resulting freed electrons register hits

along the wire. CSC chambers, spanning the endcaps of the CMS detector, consist of positively charged anode wires crossed with negatively charged copper cathode strips within a gas volume. When the gas is ionized by a muon, the electrons drift to the anode wires, while the positive ions move toward the cathode strips. This process is depicted in Figure 6.4.

The CMS muon system is designed to have the capability of reconstructing the momentum and charge of muons over the entire kinematic range of the LHC. Due to the shape of the magnet, the muon system naturally has a cylindrically shaped barrel section and two planar endcap regions. The DT chambers cover the pseudorapidity region $|\eta| < 1.2$ and are organized into 4 stations. Each of the first 3 stations contain two groups of 4 chambers. The first group measures the muon position in the r-ϕ plane, while the second group of chambers measures the position in the z direction. The two groups of stations are well separated to achieve good angular resolution. The number of muon chambers per station and the corresponding orientation was chosen such that muon hits from different stations could be optimally linked to provide good reconstructing efficiency for real muons from hard scatter events while also maintaining high background rejection. Figure 6.5 shows a cross-sectional view of the muon DT barrel system. The CSC's in the endcap regions cover the pseudorapidity region betwen $|\eta|$ values of 0.9 and 2.4. There are 4 stations of CSC's in each endcap, positioned perpendicular to the beam line, and running radially outward providing precision measurements of the position in the r-ϕ plane.

As will be shown in the sections to follow, muon reconstruction efficiency (using only information from the muon system - standalone muon) is typically $> 80\%$ except in the region $|\eta| = 1.2$, where the efficiency significantly drops due to the transition region between the DT and CSC systems. Furthermore, the reconstructed standalone muon momentum resolution is $< 10\%$ for small values of momentum, and between

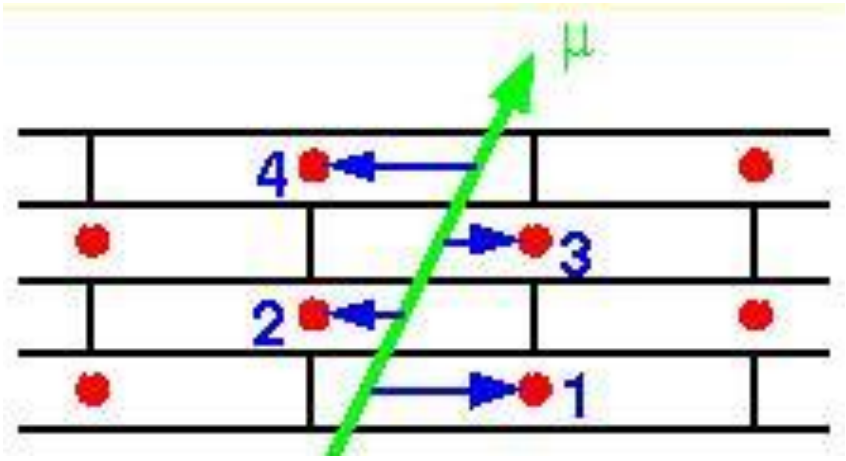

Fig. 6.4. Depiction of a Muon traversing the CSC's, creating a current, a giving rise to muon "hits."

15% and 40% at 1 TeV. However, the resolution significantly increases, even at 1 TeV, when information from the muon system is combined with the information from the inner tracker.

6.5 Calorimeters

Calorimeters are devices that measure the energy of impinging particles. Sufficiently energetic electrons and photons initiate a cascade of processes called electromagnetic showers, while charged pions, protons and neutrons initiate hadronic showers.

Electromagnetic showers are created when highly energetic photons or electrons interact in the calorimeter material and produce a cascade of electrons or photons.

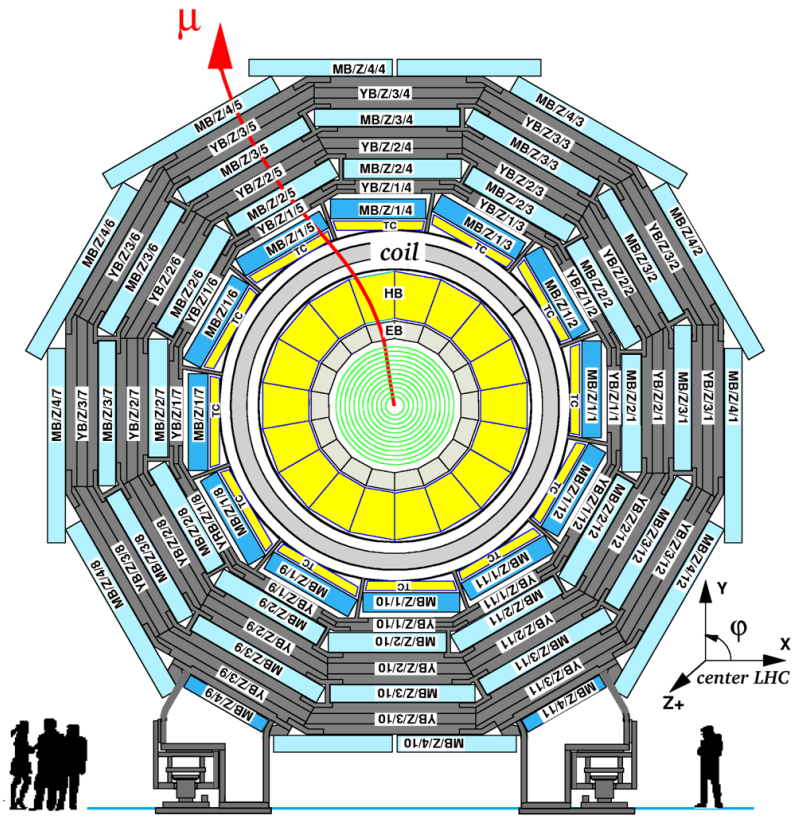

Fig. 6.5. Cross-sectional View of the Muon DT Barrel System.

If a highly energetic photon interacts in the calorimeter material, it can *convert* into electron-positron pairs. Similarly, electrons traversing the calorimeter material will decelerate and emit bremsstrahlung photons. The rate and energies by which these bremsstrahlung photons are produced are determined by the rate of deceleration according to the Bethe Bloch equation [10]:

$$\frac{dE}{dx} = -k\frac{Z}{A}\frac{\rho}{\beta^2}[ln\frac{2m_e c^2\beta^2 E_M}{I_0^2(1-\beta^2)} - 2\beta^2] \tag{6.4}$$

Electromagnetic showers are often described in terms of the radiation length X_0. A radiation length is defined as the distance a particle must travel through the calorimeter material so that the energy loss to photons is such that the particle energy decreases by a factor of $1/e$. For materials made out of only one atom, the radiation length can be approximated by:

$$X_0 \sim 180\frac{A}{Z^2} \tag{6.5}$$

where A is the atomic weight and Z is the atomic number. Calorimeters are designed such that the thickness of the calorimeters contain enough radiation lengths to accurately measure electron and photon energies. The process by which an electromagnetic shower is produced is depicted in Figure 6.6. The initial photon, with energy E_0 undergoes electron-positron pair creation with each particle carrying an energy of $E_0/2$. After another radiation length, the electron and positron each emit bremsstrahlung photons therefore resulting in four particles each with energy approximately $E_0/4$. This *cascade* process continues until the particles have low enough

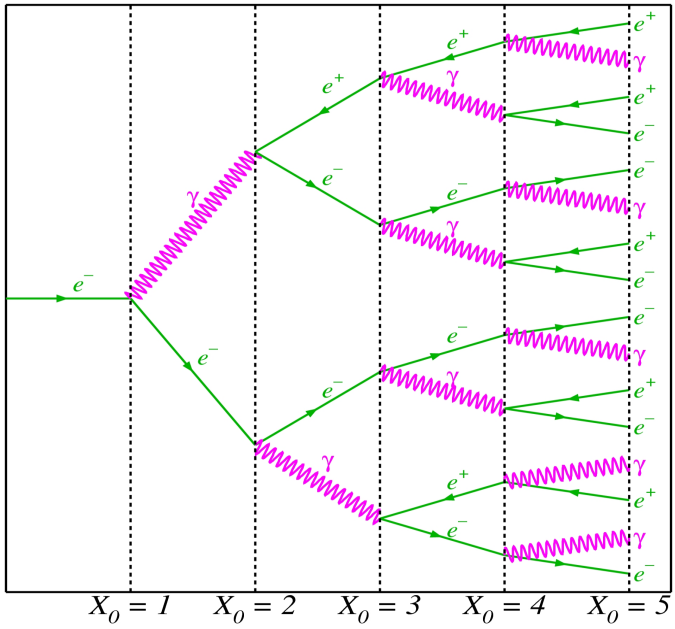

Fig. 6.6. Depiction of the development of an electromagnetic shower [11].

energies that other effects dominate.

Knowledge of the transverse and longitudinal profiles of showers produced by electrons and photons is required to accurately determine the initial energy of the electron or photon. There are many models used to predict the shower profiles, however, the simplest model assumes the depth of the electromagnetic shower is dependent on the energy of the incident particle. A depth of approximately $20X_0$ is required to contain a shower from a $\sim$ 50 GeV electron or photon. The transverse shower profile of incident particles has a much smaller dependence on the energy of the incident particle. The dependence can be quantified in terms of the Moliere radius and defined as [12]:

$$R_M = \frac{21.2 \text{ MeV}}{E_c} X_0 \tag{6.6}$$

where E_c is the energy below which the particles can no longer cascade. E_c is approximately 100 MeV [12]. An electromagnetic shower will usually deposit 95% of its energy within two R_M of the shower axis.

Hadrons also create calorimeter showers; however, they are produced through the strong interactions between the impinging particles and atomic nuclei. Since hadrons are more massive than electrons, they do not produce calorimeter showers associated with the ionization of the molecules through emission of electrons. Additionally, the probability for such interactions to occur are much smaller than the case of electromagnetic showers from electrons and photons. For these reasons, hadronic calorimeters are designed with substantially larger amounts of material. Unlike electromagnetic showers, which are mostly described in terms of radiation lengths, hadronic showers are described in terms of absorption lengths λ [13]:

$$\lambda = \frac{A}{N_A \sigma_{abs}} \tag{6.7}$$

where A is the atomic weight, N_A is Avogadro's number 6.022 x 10^{23} atoms/mol, and σ_{abs} is an absorption cross-section defined as:

$$\sigma_{abs} = \sigma_T - \sigma_{el} - \sigma_q \tag{6.8}$$

where σ_T is the total interaction cross-section, σ_{el} is the cross-section for elastic scattering off a nucleus, and σ_q is the cross-section for quasi-elastic scattering off nucleons. For example, the longitudinal development of a hadronic shower from a 50 GeV hadron is mostly contained in 5λ [13].

6.5.1 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (Ecal) is designed to reconstruct and identify electrons and photons. It is composed of Lead Tungstate crystals that scintillate when electrons or photons interact with them. Photo-detectors are glued to the back of the crystals to detect the scintillation light and convert it to an electrical signal. When a scintillation photon strikes the silicon of the photo-detector, an avalanche of electrons is created giving rise to a current that is amplified, digitized, and then transported away by fiber optic cables. Figure 6.7 shows a depiction of an electron interacting with the Ecal crystals and creating an electromagnetic shower in the calorimeter. The Ecal is made up of a barrel section and two end-cap sections, forming a layer between the tracker and hadronic calorimeter sub-detectors. The barrel section consists of 61,200 such crystals divided into 36 modules each weighing around 3 tons and containing 1700 crystals. The end-cap sections seal off the barrel at each end and are made up of approximately 15,000 additional crystals. Finally, in order to distinguish a single very energetic photon from two less interesting low energy collinear photons, the Ecal also contains pre-shower detectors that sit in front of the endcaps. Figure 6.8 displays the layout of the CMS Ecal subdetector.

6.5.2 The Hadron Calorimeter

The Hadron Calorimeter (Hcal) is a sampling calorimeter that measures the energy of hadrons. It is composed of alternating layers of absorber material (brass or steel) and tiles of plastic scintillator. When a hadron interacts with the absorber material, a cascade of particles develops. The cascade of particles then interact with the scintillator and produce blue-violet light. Optical fibers then absorb this light and shift the the wavelength into the green light region so the optical cables can carry the light away to readout boxes (RBX). The signals from successive tiles are

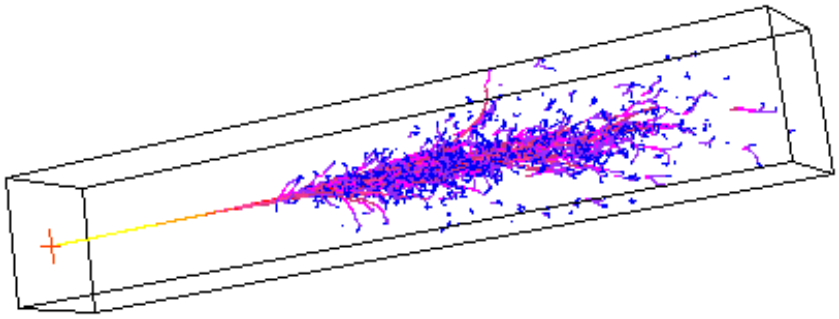

Fig. 6.7. Electron interacting with the Ecal crystals and creating an electromagnetic shower in the calorimeter.

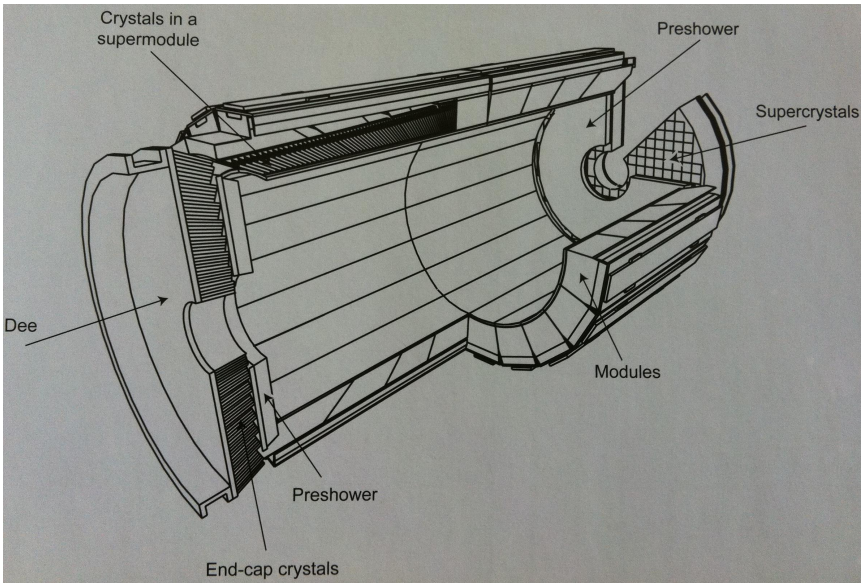

Fig. 6.8. View of the Ecal Sub-dectector.

then combined to form towers. The final combined optical signals are converted into electronic signals by Hybrid Photo-diodes (HPD) via the photoelectric effect.

6.6 Data Acquisition and Triggering

At design luminosities, the LHC is expected to produce approximately one billion *pp* collisions per second. Additionally, each event/collision can produce approximately 1000 particles. Therefore, it is not possible to store the information from every single collision. Furthermore, even if it was feasible to store the information from every single event, it is highly probable that most of those events will not contain physics signatures of interest for searches beyond the SM.

A *trigger* system is used to select "interesting" events and reduce the event rate to a manageable quantity, approximately 100 Hz. Because groups of protons are colliding in bunches separated by 25 nanoseconds, a new batch of particles are being produced before the previous batch of particles have traversed the CMS detector. This problem is tackled by storing data in pipelines that can simultaneously retain information from several collisions. The good time resolution of the detector and the synchronization of the electrical readout modules ensure particles from different collisions are identified and associated to the correct event.

The first round of triggering, Level 1 (L1), is a completely hardware based process that makes decisions based on very simple considerations such as energy. The hardware based trigger process is designed to reduce the rate from 10^9 Hz to approximately 10^5 Hz. Once a manageable rate is achieved, the next round of triggering, High Level Trigger (HLT), is a software based process that can use more complex algorithms to select interesting events. For example, Ecal energy deposits can be used in combination with the tracker tracks to trigger on electron-like objects. The HLT is designed to reduce the event rate from 10^5 Hz after L1 to 10^2 Hz to be used

for the final analysis. Because the luminosities can increase throughout any given data taking period, the trigger requirements can change often in order to maintain a manageable rate, making it difficult to emulate the trigger in the simulated samples. The details related to the specific triggers that are used in this analysis are described in section 14.

7. TAU LEPTONS

7.1 Properties of Taus

Taus are the heaviest leptons currently known, with a mass of 1.777 GeV/c^2 and a lifetime of 2.9 x 10^{-13} seconds. It is the only lepton with enough mass to decay both leptonically and hadronically. The decay channels always involve a mediator W boson and a tau (anti) neutrino. About one third of the W bosons decay into lighter leptons plus another (anti) neutrino, and the remainder into hadrons. In the latter case, taus consist of final states with one or three charged mesons (exceptionally more), and a few neutral pions. Taus that decay to hadrons are often referred to as *tau jets*. Table 7.1 lists the most significant tau decay modes as well as the branching ratios for those decays [14], while Figure 7.1 displays the Feynman diagram for leptonic tau decays and Figure 7.2 displays the Feynman diagram of a possible hadronic tau decay mode. Taus decay to hadrons approximately 65 % of the time and to lighter leptons approximately 35 % of the time.

Having a lifetime of $ct \sim 87\mu m$, tau leptons decay almost instantaneously after they are produced and thus it is not possible to distinguish leptonic tau decays from the direct production of electrons and muons. Therefore, tau identification consists of identifying taus in cases where it decays to final states consisting of hadrons (τ_h). Although the lifetime is short enough that leptonic decays cannot be distinguished from direct production of electrons and muons, it is not long enough to traverse the CMS detector material before decaying. Therefore, the presence of hadronically decaying tau leptons must be inferred from identifying their characteristic tau decay products in the detector.

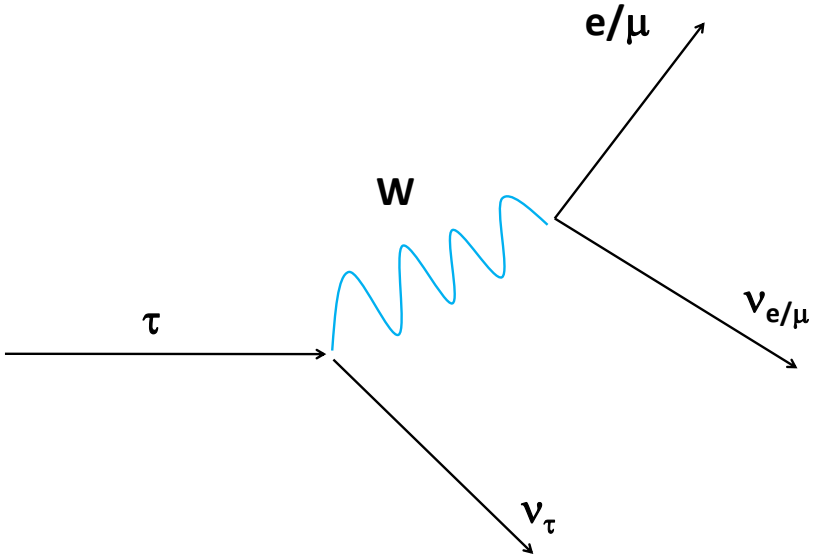

Fig. 7.1. Feynman diagram of leptonic tau decays.

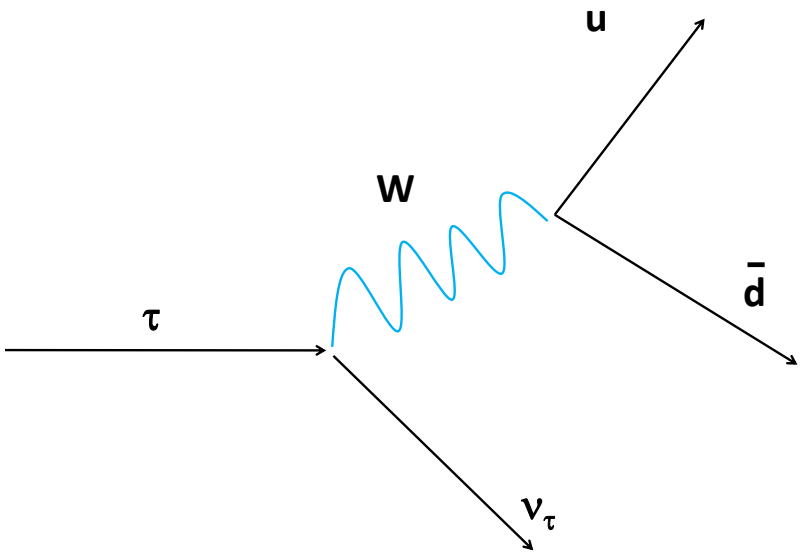

Fig. 7.2. Feynman diagram of a possible hadronic tau decay mode.

Table 7.1
τ Decay Modes [14]

Decay Mode	Branching Fraction (%)
$\tau^{\pm} \to e^{\pm}\nu_e\nu_\tau$	17.8
$\tau^{\pm} \to \mu^{\pm}\nu_\mu\nu_\tau$	17.4
$\tau^{\pm} \to \pi^{\pm}\nu_\tau$	11.1
$\tau^{\pm} \to \pi^{\pm}\pi^0\nu_\tau$	25.4
$\tau^{\pm} \to \pi^{\pm}\pi^0\pi^0\nu_\tau$	9.2
$\tau^{\pm} \to \pi^{\pm}\pi^0\pi^0\pi^0\nu_\tau$	1.1
$\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}\nu_\tau$	9.5
$\tau^{\pm} \to \pi^{\pm}\pi^{\pm}\pi^{\mp}\pi^0\nu_\tau$	4.4

7.2 Experimental Challenges for Taus at Hadron Colliders

In section 13.4, the details of tau reconstruction from the detector point of view will be described. However, in general, the main challenge in identifying taus at hadron colliders is that they closely resemble the physics signature of generic quark/gluon QCD jets (defined in the next section) that are produced with several orders of magnitude larger than any new physics signatures. Furthermore, because tau leptons decay to neutrinos which escape the detector undetected, a significant fraction of the tau momentum is lost, rendering hadronic tau jets even softer and further reducing the possibility of successful discrimination from the many quark/gluon related backgrounds. In searches for new physics using $\tau\tau$ final states, an additional complication arises due to the inability to reconstruct a narrow $\tau\tau$ mass resonance due to the loss of energy from the neutrinos.

7.3 Why Taus?

The LHC is expected to produce copious amounts of W and Z bosons that can decay tau leptons. Therefore, physics with tau leptons provide the possibility of calibrating and understanding the various systematic effects arising from the imper-

fections of the CMS detector. Furthermore, SM processes with tau leptons can be used to measure important quantities such as τ identification efficiencies for physics searches beyond the SM. For example, the SM $Z \rightarrow \tau\tau$ process will be the main channel of interest in early data due to the presence of an additional tau lepton that can be used to further suppress backgrounds. These studies usually consider the case where one tau lepton decays hadronically while the other decays leptonically. This also allows one to select events using much simpler electron or muon triggers, leading to unbiased samples of τ_h candidates. This characteristic will be exploited in this analysis as a means of determining the level of consistency between simulation and data and the extraction of the τ identification efficiencies and scale factors.

Tau leptons can become the preferred final states for new and exciting physics models such as the Higgs boson, minimal extensions of the standard model, and Supersymmetry (SUSY). SUSY models with R-parity invariance such as minimal supergravity (mSUGRA) naturally give rise to a cold dark matter candidate (CDM), generally the lightest neutralino ($\tilde{\chi}_1^0$). Additionally, cases with large tanβ give rise to the correct mass hierarchy needed so the dark matter candidates could interact in the early Big Bang "just right" to produce the correct CDM relic density observed today by WMAP [1]. The region where the correct dark matter allowed mass hierarchy is achieved is called the stau-neutralino coannihilation region. This region has a striking characteristic that the $\tilde{\tau}_1$ and $\tilde{\chi}_1^0$ are nearly mass degenerate, $\Delta M = M_{\tilde{\tau}_1} - M_{\tilde{\chi}_1^0} \sim 5 - 15$ GeV, leading to final states involving mostly tau leptons. Moreover, this region could provide the only experimental allowed means by which the correct dark matter relic density can be achieved. In [15, 16], it is shown that final states with pairs of tau leptons can be used to search for SUSY and provide precision measurements of the SUSY masses and the cold dark matter relic density.

In minimal supersymmetric extensions of the SM (MSSM), the presence of two Higgs doublets leads to a more complicated Higgs boson sector, with five massive Higgs bosons: a light neutral scalar (h), two charged scalars ($H^{\pm}$), a heavy neutral CP-even state (H) and a neutral CP-odd state (A). The masses of the Higgs bosons mainly depend on two parameters, the mass of the pseudoscalar state m_A, and tanβ. In a large part of the parameter space, the branching ratio to pairs of tau leptons is approximately 10 %. In these cases, a search for these Higgs bosons decaying to taus is the best experimental signature given that the decay mode with b quarks, although produced with a larger branching fraction, suffers from overwhelming background and mass resolution effects.

8. HADRONIC JETS

One of the main experimental challenges with identifying hadronically decaying tau leptons is they resemble the detector signature of quark or gluons that produce hadrons. The natural question is how two seemingly different particles can produce similar physics signatures in the CMS detector. This section describes the *hadronization* of partons responsible for the difficulty in discriminating τ_h's from quark and gluon signatures.

Although protons are believed to be composed of u and d quarks, experimental observation has shown protons are likely to have a much more complex makeup when probed at such high energies as is the case at the LHC. Protons are held together by a complex network of virtual gluons, virtual quarks, and processes that help maintain the bounded state of the bare quarks. These virtual gluons (which mediate the strong force), virtual quarks, and bare quarks are collectively known as *partons*. When two protons collide at the LHC, it is the interactions between these partons that generate the hard scatter event responsible for the SM processes such as Z boson production and BSM processes such as sequential Z' production. In fact, as will be discussed in section 9, the momentum of the proton is divided among the constituent partons. Therefore, when probing for new physics signature, each event is sensitive up to the energies of the partons involved in the hard scatter event (as opposed to the energy of the colliding protons). The advantage of this characteristic is that one can search for a broad energy spectrum of new physics signatures. However, the disadvantage is that protons are essentially "ripped" apart when they collide, leading to a large number of hadrons which create structures called *jets* that mimic the detector signature of hadronically decaying tau leptons.

When two partons from the incoming protons interact, they often produce pairs of gluons and quarks traversing in opposite directions (in the center of mass frame of reference). However, because the strong force does not allow for asymptotic freedom, a process known as *hadronization* [17] causes the color field of these seemingly free partons to break into two through the creation of quark-antiquark pairs. The initial seemingly free partons contain some color field mediated by gluons. Because the strength of the strong force is directly proportional to the distance between these partons, as the partons continue to move away from each other the color field becomes strong enough to produce additional quark-antiquark pairs. At this point, the final state consists of four partons all mediated by gluons producing two colorless bound states traveling in opposite directions. This process of hadronization can be visualized by making an analogy to an elastic rubber band that is pulled apart from both ends until it breaks into two bands. The ends of the elastic band represent quarks, while the band itself represents the gluon field.

The hadronization process will continue to produce quark-antiquark pairs as long as the gluon fields have sufficient energy to produce new particles. When the gluon fields no longer have sufficient energy, the quarks and anti-quarks that remain will combine to form bound states of particles called Hadrons. The process of hadronization is illustrated in Figure 8.1. Therefore, the end result of these typical hard scatter events is two collimated sprays of hadrons which are referred to as *hadronic jets*, or *QCD jets*. Because these hadronic jets are a result of quark-antiquark pairs producing bounded states, jets are usually composed of one of the most common hadrons, *pions*. Because τ leptons can decay to final states containing mostly one or three charged pions and a few neutral pions, at first glance they resemble the makeup of a QCD jet.

Hadronic jets are produced in large quantities at the LHC. For example, QCD dijet production has a cross-section that can be approximately 10^6 larger than pseu-

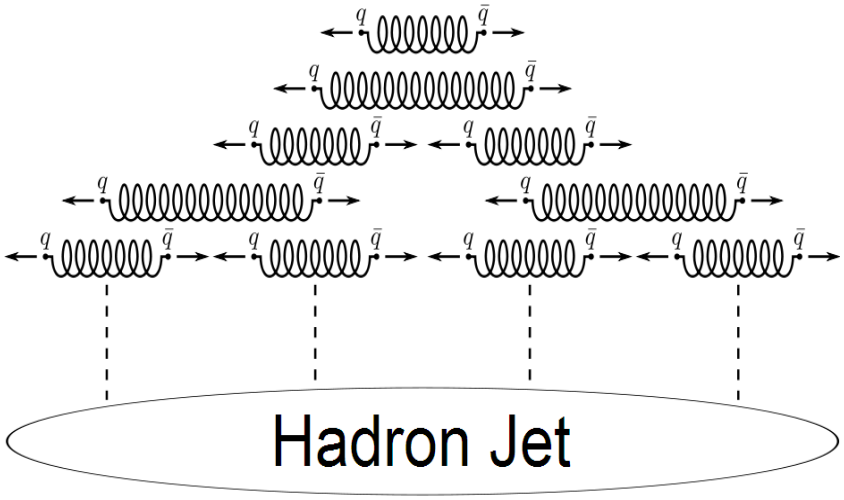

Fig. 8.1. Depiction of jet hadronization.

doscalar Higgs production or sequential Z' production. Therefore, a search for new massive resonances decaying to pairs of τ leptons must be designed with good jet reduction as the main focus.

Fortunately, hadronic jets are different from hadronic τ decays in a few key aspects, making τ_h identification possible. For example, the transverse width of hadronic jets is a direct result of the energy stored in the gluon fields of the fragmenting hadrons, whereas the transverse width of τ_h's is completely driven by the kinematics of a τ decaying to the constituent final state hadrons and neutrinos. Therefore, hadronic jets tend to have a wider energy spread or energy profile than hadronically decaying tau leptons. Taus are often characterized as having a "pencil like" structure due to the narrow collimated final state of a few energetic particles. The "pencil like" structure of the τ_h is depicted in Figure 8.2. Therefore, as will be described in section 13.4, τ_h identification consists of defining a nar-

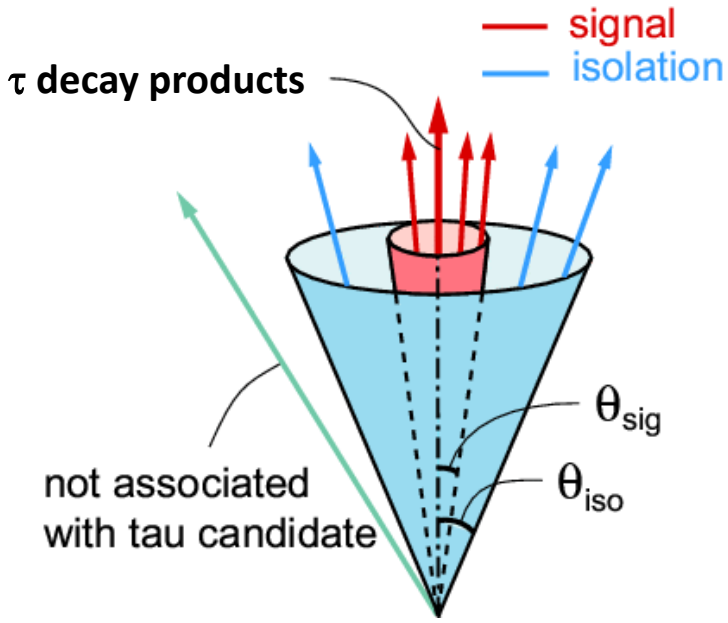

Fig. 8.2. Sketch depicting the narrow "pencil like" characteristic of hadronic τ decays [3].

row region around the visible τ direction to contain the τ decay products and a larger region to quantify the width of the energy profile of hadronic jets. Figure 8.3 shows the number density of charged hadrons as a function of the distance $\Delta R = \sqrt{(\eta_{\tau/jet} - \eta_{\text{hadron}})^2 + (\phi_{\tau/jet} - \phi_{\text{hadron}})^2}$ between the τ/jet direction and the charged hadron direction. It is clear that hadronic jets (open black circles) tend to a larger multiplicity of particles and a wider density profile. Similary, Figure 8.4 shows the number of charged particles (*prongs*) within a narrow region around the τ direction. As expected, taus are characterized by a low multiplicity of charged particles, while jets are characterized by a larger multiplicity of charged particles. The details of these results will be explained in section 13.4.

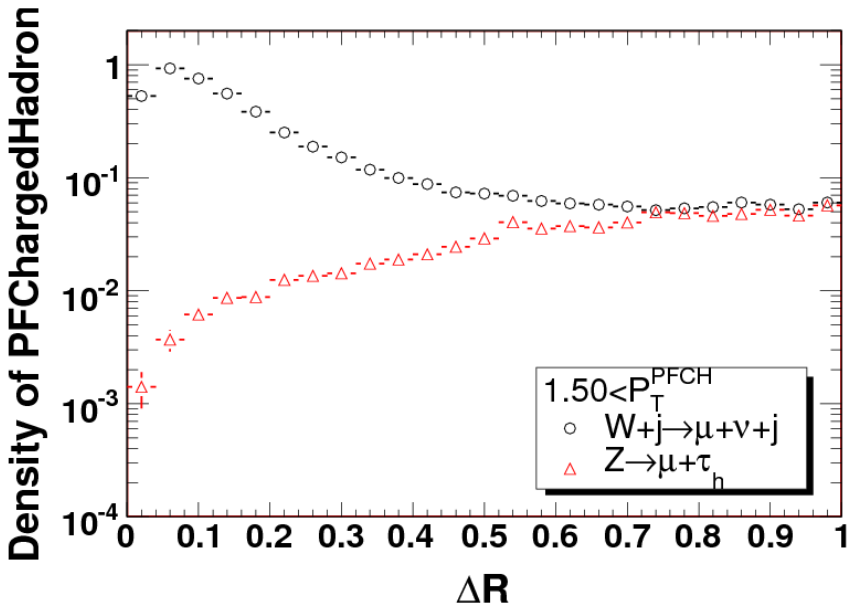

Fig. 8.3. Density of PF charged hadrons with respect to the τ_h (red open triangles) and quark/gluon jet (black open circles) directions.

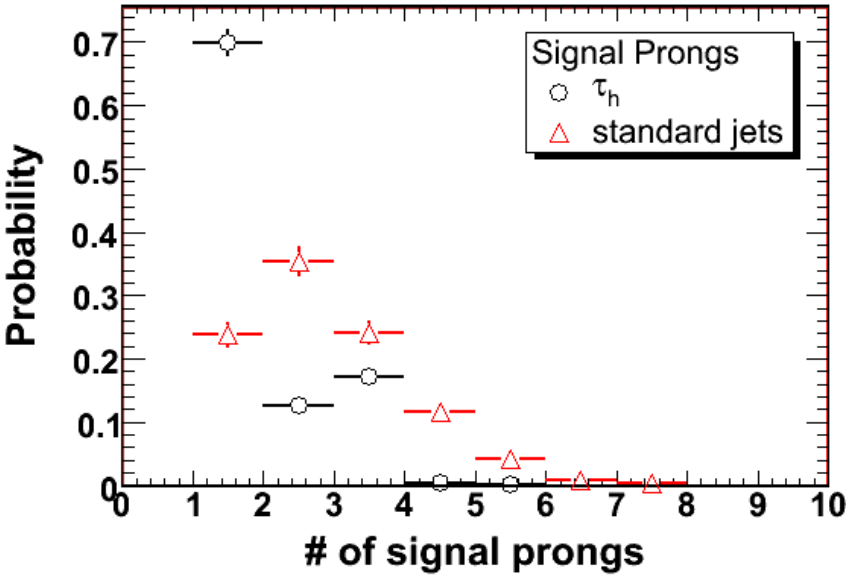

Fig. 8.4. Number of PF charged hadrons in a narrow region around the τ_h (black open circles) and quark/gluon jet (red open triangles) directions.

9. SIMULATION OF EVENTS

Designing an analysis that is robust and achieves the desired sensitivity is difficult to do if one merely considers the theoretical equations that drive the underlying physics of interest. In high energy physics, which at the moment is driven by the SM, one can never know with 100% certainty whether a given outcome *will* occur. In many cases, the only statement that can be made about a physical system is the it probability for which a certain outcome *can* occur. For example, in the context of hadron colliders, if a given number of events N are produced, it cannot be determined with 100% certainty that n of those events are due to e.g. $pp \to Z \to \tau\tau$ production. However, one can make the following statement: "Suppose there is some probability p that $pp \to Z \to \tau\tau$ production will occur. On *average*, from the N observed events, one might *expect* $n = pN$ events due to $pp \to Z \to \tau\tau$ production." Therefore, constructing an analysis requires the generation of an ensemble of simulated events that can be used to determine the characteristics of the signal process of interest on a statistical basis.

The term *Monte Carlo* (MC) is used to describe any technique that uses random numbers to solve a numerical problem. Monte Carlo techniques are often used to simulate or solve complex physical or mathematical systems when obtaining the exact result using deterministic equations or algorithms is not possible. In statistical analysis, MC methods are employed when the model of interest contains significant uncertainty for the inputs. In mathematics, MC methods can be utilized to solve complex integration problems. In fact, MC integration methods will be used to extract the statistical significance of the results presented.

At CMS, the generation of events is split into two main steps. The first step, performed by an *event generator*, uses a model based on the SM or a BSM theory to

generate the particles believed to be produced during pp collisions. The second step then simulates the way those particles interact with the CMS detector and how that translates to the electrical signals that are used to reconstruct an event. The MC samples of generated events and simulated interactions with the detector are used to guide the design of an analysis and make comparisons with the data.

In pp collisions at the LHC, the production of other particles occur mostly because of the interaction of the partons within the protons. Therefore, the first step in the generation of MC samples is to calculate the parameters relevant to the creation of other particles from the incoming partons. To carry out these non-trivial calculations, the energy of each parton must be used as an input. However, because protons are composed of *up* and *down* quarks joined together by a "sea" of gluons, the initial parton energy is not known. If the proton is traveling with some energy E_p, each parton within the proton can have an energy of xE_p where the factor x varies between 0 and 1. The value of x is described by the parton distribution functions (PDF), which have been measured based on many experiments such as ZEUS, H1 experiments on the HERA accelerator at the DESY laboratory, and CDF and D0 experiments at the Fermi National Accelerator Laboratory (FNAL). In fact, events from collisions at the LHC will be used to reduce the uncertainty on the the PDF's and provide better calculations of important quantities such as production cross-sections.

At the theoretical level, it is fairly straight forward to calculate matrix elements with "decent" accuracy (leading order calculations, LO). The level of accuracy is dependent on the analysis of interest. Depending on the scope of the experiment and a particular analysis, it can become difficult and cumbersome to achieve accuracy beyond LO calculations. For example, it has been shown that the LO diagram for the production of a SM Z boson occurs via $q\bar{q}$ annihilation as depicted in Figure 4.1. However, in the next to leading order (NLO) matrix elements, additional effects

such as the exchange of gluons between initial state quarks also exist. Therefore, the NLO diagrams can take place via $qg \rightarrow Zq$ or $qq \rightarrow Zg$. The number of Feynman diagrams and required mathematical calculations increase significantly when one considers higher order processes. In collider experiments, the perfect calculation of the full matrix elements are not required. The only requirement is that events are generated with "good enough" accuracy that our MC samples allow us to study the characteristics of a particular process so the design of a robust analysis is feasible. For example, in the case of the production of a Z boson, one is well aware that the final state will always be Z+n jets. Once we know this, then the important question is how one can describe the n jets without the explicit calculation of the full blown matrix element. *Parton showering* is the process by which these effects are modeled.

Parton showering is built upon the physically well motivated belief that the radiation from partons can be built up from a series of similar processes where a parton losses energy as it radiates a gluon ($q \rightarrow qg$). The *shower* begins at the energy scale of the hard process or initial interacting partons and continues to evolve until the radiated partons do not have sufficient energy to continue to shower. Parton showering is described by the DGLAP equations [18]. Because the strong force does not allow for asymptotic freedom, the partons produced through the showering process must undergo the simulation of the hadronization process. Hadronization is calculated using non-perturbative models.

After the showering and hadronization of partons, the MC process consists of simulating the decay of all remaining unstable particles such as τ leptons. These decays are usually performed using purpose built software routines that contain all the necessary information such as the possible decay modes and branching ratios. In the case of tau leptons, the TAUOLA [19] package is used. The TAUOLA package not only provides the correct branching ratios for tau decays, but also accounts for

the polarization of tau leptons that affect the momentum distribution of the corresponding decay products.

The PYTHIA [20] event generator is a leading order (LO) generator, which means the underlying physics processes are produced using only tree level Feynman diagrams. General purpose generators, such as PYTHIA, are capable of performing all of the MC steps described above. MADGRAPH [21] allows the generation of additional hard parton radiation and can be interfaced with general purpose generators such as PYTHIA to perform the showering and hadronization steps.

Interactions of final state particles with the CMS detector are simulated with GEANT4 [22]. GEANT4 contains information such as the detector geometry and material budget needed to accurately model electromagnetic showers, hadronic showers, and the trajectory of charged particles in the magnetic field.

The last stage of the generation of MC samples involves the proper modeling of pile-up interactions. At design LHC luminosities, an average of 20 secondary pp interactions of lower energy can be superimposed on the hard scattering event of interest. These PU interactions can have significant affects on important physics quantities such as the reconstruction of the momentum imbalance in the detector. Therefore, very careful attention must be placed to either model the correct PU contribution or apply proper corrections to the MC sample. There are two methods to incorporate PU effects at CMS. The first method uses real CMS events overlaid on top of events from the MC generation. At CMS, this process is referred to as *data mixing*. The events "mixed" into the MC generation are required to pass special triggers designed to select low bias events and be timed in with the bunch crossings of the LHC. Modeling of PU using this method requires the MC to have the same distribution of PU interaction vertices with that observed in data. However, this

is often difficult to do. Therefore, a second procedure is often used. MC events can be generated with one of the following conditions: (1) uniform distribution of PU vertices; (2) with no contribution from PU interactions. In each scenario, the MC can then be reweighted so as to achieve the correct distribution of PU vertices. This method has the advantage that MC samples do not have to be re-generated as the instantaneous luminosity increases. Furthermore, the reweighing technique allows one to reweigh events based on the physics of interest. For example, it is not generally true that $t\bar{t}$ events will contain a similar distribution of vertices as events arising from Z production. $t\bar{t}$ events very often contain additional secondary vertices due to the presence of long lived particles such as b quarks, whereas e.g. $Z \to ll/q\bar{q}$ events very rarely contain additional vertices. Therefore, the method of reweighing MC events to account for PU effects allows one to determine the proper contribution and distribution of PU based on control samples derived directly from data.

10. ANALYSIS STRATEGY

As the tau lepton decays to $e\overline{\nu}_e\nu_\tau$ (17.8%), $\mu\overline{\nu}_\mu\nu_\tau$ (17.4%), and hadrons + ν_τ (64.8%) there are six possible analyses for pairs of tau leptons: (1) $\mu\mu \sim$ (3.1%), (2) $ee \sim$ (3.1%), (3) $e\mu \sim$ (6.2%), (4) $e\tau_h \sim$ (23.1%), (5) $\mu\tau_h \sim$ (22.5%), (6) $\tau_h\tau_h \sim$ (42%). The six possible final states and their branching ratios are listed in Table 10.1. The ee and $\mu\mu$ channels are not considered in this analysis due to the small branching ratio and copious Drell-Yan $Z/\gamma^* \rightarrow e^+e^-, \mu^+\mu^-$ production. Furthermore, the ee and $\mu\mu$ final states will be difficult to distinguish from direct production of $ee/\mu\mu$ pairs from new resonances of similar mass. Choosing between the remaining four final states requires a compromise between the branching ratios and the probabilities for hadronic jets to fake the leptons. The $e\mu$ final state has the smallest branching fraction, but provides the lowest possible background contamination (jet$\rightarrow$electron fake rate is $\sim 10^{-2}$ and jet$\rightarrow \mu$ fake rate is $\sim 10^{-4}$–10^{-3}). The $\tau_h\tau_h$ final state has the largest branching fraction, but provides the largest possible background contamination (jet$\rightarrow \tau_h \sim 10^{-2}$–$10^{-1}$). Therefore, as a middle ground between branching ratios and the probabilities for hadronic jets to fake leptons, this thesis mostly focuses on a search for new massive resonances decaying to pairs of τ leptons $X \rightarrow \tau\tau$ using the final state in which one τ decays leptonically to a μ, while the other τ decays to hadrons.

In general, events are selected with two oppositely charged, nearly back–to–back objects. Because τ leptons decay to neutrinos via a virtual W boson, the $\tau\tau$ system can have up to four neutrinos. In the case of the $\mu\tau$ final state, there are three neutrinos that escape the detector undetected. Therefore, events are expected to have a momentum imbalance in the detector that can be exploited to discriminate from background processes. In contrast to searches in the ee and $\mu\mu$ channels, the visible $\tau\tau$ invariant mass does not produce a narrow peak due to the missing neu-

Table 10.1
$\tau\tau$ Branching Ratios.

Final State	Branching Ratio (%)
$\tau\tau \to \mu\mu$	3.1
$\tau\tau \to ee$	3.1
$\tau\tau \to e\mu$	6.2
$\tau\tau \to e\tau_h$	23.1
$\tau\tau \to \mu\tau_h$	22.5
$\tau\tau \to \tau_h\tau_h$	42.0

trinos. However, the momentum imbalance in the detector can be used to calculate the mass of the $\tau\tau$ system and provide better separation between signal and background. However, as will be seen in the sections that follow, regardless of how one calculates the $\tau\tau$ mass, a narrow mass resonance (when compared to ee and $\mu\mu$ mass resolutions) cannot be achieved. Instead, a search is made for a broad enhancement of the $\tau\tau$ invariant mass distribution consistent with new massive resonance production and incompatible with SM expectations. The statistical extraction of the signal significance depends on the signal to background rejection as well as the systematic effects from the lack of knowledge of various efficiencies, shapes, and theoretical considerations. Therefore, in designing and selecting the selections, significant attention is placed on using criteria that maintains high efficiency for signal events, provide strong background suppression, and reduce the influence of systematic effects.

Since Drell-Yan processes $pp \to \gamma^*/Z \to \tau\tau$ are one of the main SM τ pair production mechanisms at the LHC, $Z \to \tau\tau$ is one of the main backgrounds. Additionally, because $Z \to \tau\tau$ mimics the final state of new massive resonances decaying to pairs of τ leptons $X \to \tau\tau$, $Z \to \tau\tau$ will be used as an important validation of the final signal selections and that the identification of τ's is indeed possible. Since a clean sample of high energy τ leptons cannot be obtained, one of the main concerns in any search for physics with high energy τ's is whether the τ identification criteria

is still effective in the particular energy regime. In fact, this is not specific to τ's and is a consideration for any analysis that deals with high p_T jets or leptons. In this analysis, this concern is tackled by (1) verifying that the τ identification criteria works at $p_T < 100$ GeV/c and is consistent with MC expectation by using data–driven techniques to obtain a clean sample of $Z \to \tau\tau$ events, (2) τ identification is well understood in the MC simulated samples, and (3) using τ identification criteria that is highly unlikely to produce differences between MC and data.

In order to ensure robustness of the analysis and the confidence in the results, whenever possible the analysis relies on the data itself to understand and validate the efficiency of reconstruction methods as well as the estimation of the background contributions. For that purpose control regions are defined with most of the selections similar to what those used in the main search but enriched with events from background processes. Once a background enhanced region is created, selection efficiencies are measured in those regions and used to extrapolate to the region where new massive resonance signals are expected. In cases where a complete data–driven method is not possible the analysis makes use of scale factors, ratio between observed data events and expected MC events, in the background enhanced region to estimate the background contribution in the signal region.

To quantify the significance of any possible excess or set upper limits on the production rate, a fit of the reconstructed $\tau\tau$ mass distribution is performed and a Bayesian technique is employed to interpret the results in terms of the upper 95% credibility level limits. The posterior probability density function (likelihood) is calculated by taking into account systematic uncertainties and the correlations of systematic effects.

11. BACKGROUNDS

The main background for this analysis is due to the irreducible Drell-Yan $pp \rightarrow \gamma^*/Z \rightarrow \tau\tau$. The $Z \rightarrow \tau\tau$ final state mimics that of a hypothetical new massive gauge boson that produces pairs of τ's. However, the key feature that distinguishes $Z \rightarrow \tau\tau$ from new hypothetical massive (few hundred GeV/c^2 to 1 TeV) $X \rightarrow \tau\tau$ production is that the mass of the $\tau\tau$ system will be located at values < 100 GeV/c^2, whereas new massive resonance production is expected to produce $\tau\tau$ pairs with mass > 100 GeV/c^2. This key feature is utilized as a means of validating the robustness of the analysis and the effectiveness of the τ identification criteria (section 17.5). The Feynman diagram for Drell-Yan $pp \rightarrow \gamma^*/Z \rightarrow \tau\tau$ production is shown in Figure 4.1.

In most analyses with hadronically decaying τ leptons, QCD dijet production is the main source of background. As discussed in section 8, hadronic jets are combinations of hadrons. For this reason, hadronic jets have non-negligible probabilities to fake τ's. Furthermore, even if object identification criteria is significantly small, QCD dijet background rates can still be non-negligible because they are produced with cross-sections that are significantly larger than the theoretical cross-sections for new heavy mass resonances. However, the advantage of using the $\mu\tau_h$ final state is that muons have the lowest jet misidentification among leptons; therefore, providing the best possible dijet rejection factors of all final states with at least one τ_h. Additionally, QCD dijet production does not have real intrinsic momentum imbalance due to neutrinos. Therefore, the momentum imbalance of the event can be utilized to further suppress QCD dijet backgrounds. Figure 11.1 shows an example Feynman diagram for QCD dijet production. Because b quarks can decay to muons ($b \rightarrow \nu_\mu \mu c$), QCD dijet backgrounds are mostly dominated by cases where $b\bar{b}$ pairs are produced.

W boson production in association with jets (W+jets) becomes a dominant background because a clean well reconstructed muon is produced by the W boson. Therefore, the requirement of a clean muon signature does not provide additional discrimination. Additionally, the neutrino from the W boson decay acquires an average energy of approximately $m_W/2 \sim 40$ GeV. Because the neutrino will escape the CMS detector undetected, the measurement of the momentum imbalance will be approximately 40 GeV. Therefore, a requirement on the momentum imbalance of the event does not provide significant discrimination against this background. Finally, the presence of n jets means that the contamination from W + n jets in the signal region is highly dependent on the jet$\rightarrow \tau_h$ fake rate, which is the largest amongst leptons. Therefore, reducing W+jets also requires additional topological requirements. The Feynman diagram for W+jets production is shown in Figure 11.2.

Events with top-antitop pairs ($t\bar{t}$) becomes a significant background because the W boson from the t decay produces a clean well reconstructed muon. Therefore, as is the case for W+jets production, the requirement of a clean muon signature does not provide additional discrimination. Also, $t\bar{t}$ is enhanced due to the presence of two W bosons, one from the t and one from the $\bar{t}$. Therefore, one of the W bosons can provide a clean muon signature, whereas the second W boson can decay in to a clean hadronically decaying τ lepton or quark-antiquark pairs that give rise to hadronic jets that fake the τ_h. Additionally, because t quarks decay to b quarks ($t \rightarrow bW$) and the b quarks can decay to μ's ($b \rightarrow \nu_\mu \mu c$) and also produce jets, the presence of b quarks provide additional means by which the μ or τ_h requirements can be satisfied. At the same time, the presence of b jets and other topological selections can be used to suppress this background. The Feynman diagram for $t\bar{t}$ production is shown in Figure 11.3.

Finally, Drell-Yan $pp \to \gamma^*/Z \to \mu\mu$ production is also a potential background for this analysis. $Z \to \mu\mu$ becomes a background because one of the muons from the Z boson decay produces a clean muon signature in a similar way to the muons produced from leptonic τ decays. In many experiments, the lifetime of the τ is utilized as a means of discriminating leptonic tau decays from direct production of leptons such as Drell-Yan $pp \to \gamma^*/Z \to \mu\mu$. However, at the moment this is not possible at CMS. Therefore, the background contamination from $Z \to \mu\mu$ events is entirely dependent on the probability for the second muon from the decay of the Z boson to fake a hadronically decaying tau lepton. In the sections that follow, it will be shown that the $\mu \to \tau_h$ fake rates are small enough that this background is not a major concern for this analysis. However, we do make use of a clean control sample of $Z \to \mu\mu$ events to validate the robustness of the muon identification criteria and that this background is indeed not a major concern for this analysis.

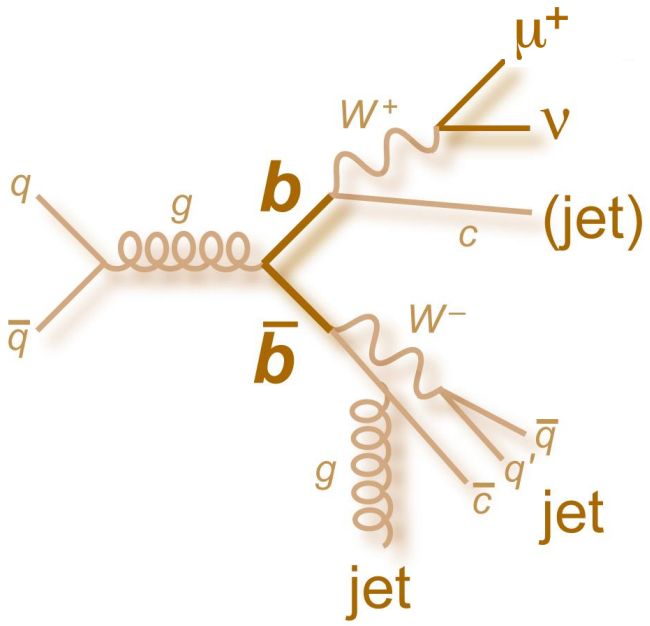

Fig. 11.1. Standard Model QCD production.

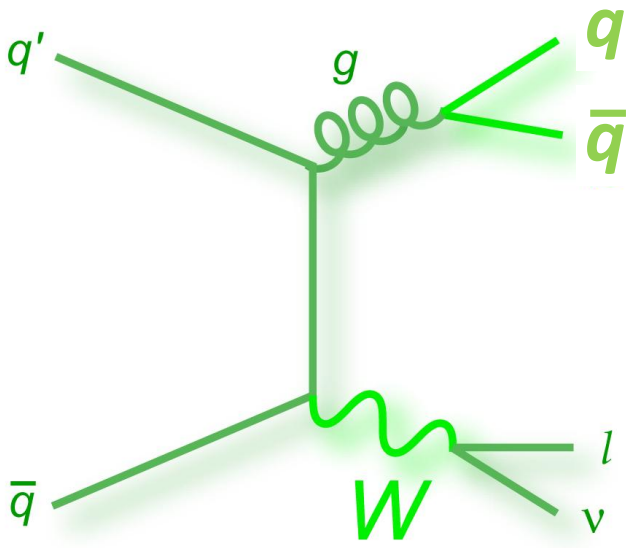

Fig. 11.2. Standard Model W+Jets production.

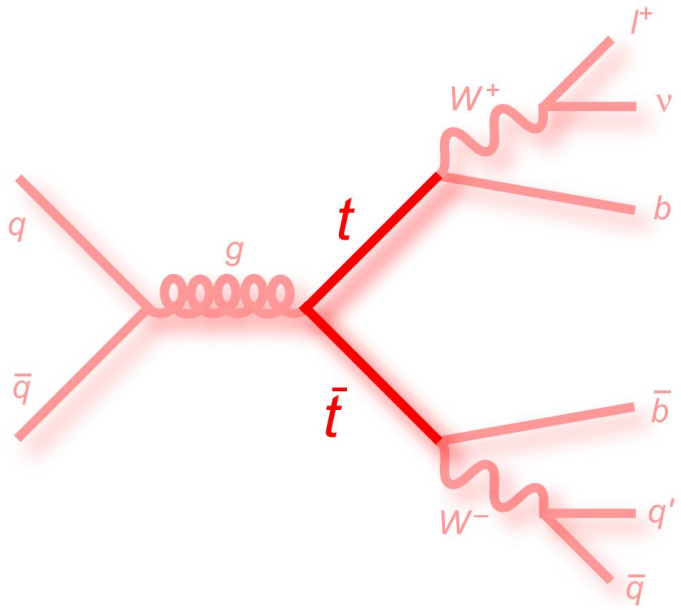

Fig. 11.3. Standard Model $t\bar{t}$ production.

12. MONTE CARLO AND DATA SAMPLES

12.1 Monte Carlo Simulated Samples

As discussed in the previous sections, the main backgrounds for the $\mu\tau_h$ final state are from $Z \to \tau\tau$, $Z \to \mu\mu$, $W + jets$, $t\bar{t}$, and QCD. Official CMS MC samples, created during Fall 2010, were produced using the MC simulation techniques and software described in section 9. The official Fall 2010 MC samples of Z^0, $t\bar{t}$, and QCD production processes were produced with the leading order event generator PYTHIA [20]. Although PYTHIA only calculates the LO cross-sections, NLO cross-sections were taken from MC@NLO [23] generators and used to scale the MC rates. However, it is important to note that although PYTHIA can only calculate the LO cross-sections, it does not present a problem for this analysis. In general, agreement between MC and data is not expected due incomplete modeling of effects such as the fragmentation of quarks/gluons, matrix element calculations, and the simulation of the amount of material budget within the CMS detector. These are just a few examples of how MC predictions can result in slightly incorrect predictions. For this reason, the MC is mainly used as a way to design a robust analysis and understand the final selection criteria. Therefore, the prediction of the background contamination in the signal regions is carried out using data-driven methods. The generation of processes leading to the production of W bosons + n jets were generated using MADGRAPH [21] in order to incorporate NLO effects that result in events with $n > 1$ jets. Because the QCD dijet production cross-sections are so large, it is not feasible to generate enough MC events expected at 36.15 pb^{-1}. Furthermore, because muons have the lowest jet misidentification among leptons, a large fraction of the QCD dijet events play no role in this particular analysis. Therefore, QCD muon enriched samples were created by using PYTHIA to simulate general QCD dijet production and selecting only those events containing at least one muon with $p_T > 10$ Gev/c. This process of selecting a subset of events is referred to as *skimming*. In

general, the only worry in using muon enriched QCD samples is the potential that the $p_T > 10$ GeV/c requirement might bias the studies. However, as will be shown in the sections that follow, the final selection thresholds are chosen well above the threshold of 10 GeV/c. Therefore, using the muon enriched sample is not expected to bias the studies. More importantly, MC and data are not expected to agree. The extraction of the QCD backgound contamination is carried out in a completely data-driven method. Therefore, although biases are not expected from using these samples, whether any bias is indeed imposed does not matter as data-driven methods are employed throughout the analysis. Finally, it is important to note that the MC samples do not contain the simulation of pile-up (PU). However, it is shown that PU effects have negligible effects on μ and τ reconstruction. Furthermore, this is validated by measuring object trigger/identification efficiencies and using control samples of $Z \to \mu\mu$ and $Z \to \tau\tau$ to show consistency between MC and data. In cases where PU effects are not negligible (e.g. measurement of the momentum imbalance), proper data-driven corrections are applied to account for discrepancies between MC and data. Table 12.1 lists the CMS MC samples used in this analysis.

Table 12.1
MC Simulated Samples.

Sample	Official CMS Datasets
$Z' \to \tau\tau$	private production (see appendix)
$Z \to \tau\tau$	*/DYToTauTau_M-20_TuneZ2_7TeV-pythia6-tauola/Fall10-START38_V12-v1/GEN-SIM-RECO*
$Z \to \mu\mu$	*/DYToMuMu_M-20_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO*
$W + jets$	*/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Fall10-START38_V12-v1/GEN-SIM-RECO*
$t\bar{t}$	*/TT_TuneZ2_7TeV-pythia6-tauola/Fall10-START38_V12-v1/GEN-SIM-RECO*
QCD	*/QCD_Pt-20_MuEnrichedPt-10_TuneZ2_7TeV-pythia6/Fall10-START38_V12-v1/GEN-SIM-RECO*

12.2 Collision Data Samples

The analysis described in this thesis is based on data taken by the CMS experiment between May 2010 and December 2010. Run 2010A and Run 2010B data for runs in the range 132440–149442 was recorded by first selecting data that was flagged as "good" by the official CMS data validation teams. Good quality data is selected by ensuring successful data acquisition from all subdetectors. Furthermore, data is flagged as either good or bad by each data quality monitoring team by ensuring that the detector components were working as expected. Runs or luminosity blocks that were tagged as bad by at least one data monitoring team are not utilized in this analysis. To ensure that the proper runs and luminosity blocks are used, the official CMS JSON file was used to select "good" run ranges and luminosity sections. With these requirements, the total validated recorded integrated luminosity used in this analysis is 36.15 pb^{-1}.

The collision data samples were reconstructed using official CMS software. As the understanding of the performance and calibration of the detector improved, new collision data samples were reconstructed to incorporate these improvements. For this analysis, the official November 4^{th} dataset is used. Table 12.2 shows the collision datasets used.

Table 12.2
Collision Data Samples.

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

13. PARTICLE IDENTIFICATION AND EVENT RECONSTRUCTION

13.1 Muon Reconstruction

Muon reconstruction is a multistep process that begins with the information gathered from the muon subdetector. The muon subdetector consists of drift tube (DT) chambers in the barrel region and cathode strip chambers (CSC's) in the endcap region. As a first step, *standalone muons* are reconstructed from hits in the individual DT and CSC chambers. A linear fit to the positions of the hits reconstructed in each of the 8-12 (in case of DT) or 6 (in case of CSC) layers of the chamber is used to construct "seeds" consisting of position and direction vectors and provide an initial estimate of the muon momentum. The seeds are then used as a starting point for the standalone fits in the muon system, which are performed using hits from the innermost muon stations combined with hits in the other muon segments using the Kalman fitting technique [24]. A suitable χ^2 cut is applied in order to reject bad hits arising from showering, delta rays and pair production. The standalone muon trajectory is reconstructed by extrapolating from the innermost muon station to the outer tracker surface. Finally, the standalone track is extrapolated to the nominal interaction point (defined by the beam-spot size: $\sigma_{xy} = 15$ μm and $\sigma_z = 5.3$ cm) and a vertex-constrained fit to the track parameters is performed.

The standalone trajectory is then used to find a matching track reconstructed in the inner silicon tracker. The best-matching tracker track is selected. The matching is performed by using the standalone reconstructed muon to extrapolate from the innermost muon station to the outer surface of the silicon tracker. Energy loss due to multiple scattering and interactions in the material are taken into account. For each tracker track-standalone muon pair, a track fit using all hits in both tracks is performed, taking into account the average expected energy losses, the magnetic field, and multiple scattering in the detector materials. The tracker track reconstruction

algorithm [25] consists of (1) the use of tracker seeds to construct initial trajectories, (2) a trajectory cleaner that resolves any ambiguities between multiple trajectories arising from the same seeds, and (3) a final fit that uses high quality hits as input to the Kalman fitting algorithm. As was the case for standalone muons, resolving any ambiguities and selecting high quality hits is done on the basis of the χ^2 of the fits. Finally, standalone muons and matching silicon tracks are used to perform a global fit resulting in a *global muon*. Figure 13.1 shows the global muon p_T resolution for signal and various backgrounds relevant to this analysis. Because the momentum resolution is expected to be better for low p_T muons where the bending of the trajectories is larger, muons from $Z \to \tau\tau$ events ($p_T^\mu \sim M_Z/6 \sim 15$ GeV/c) have the best resolution, while muons from sequential $Z' \to \tau\tau$ events ($p_T^\mu \sim M_{Z'}/6 \sim 100$ GeV/c) have the worst resolution. The p_T resolution for muons from $Z \to \tau\tau$ is $< 1\%$ and $\sim 1\%$ for $Z' \to \tau\tau$. Figure 13.2 shows the global muon reconstruction efficiency as a function of muon p_T. The reconstruction efficiency is $\sim 99\%$ for $p_T > 10$ GeV/c. Because the efficiency is a steeply falling distribution below 10 GeV/c, the analysis makes use of muons with $p_T > 20$ GeV/c in order to ensure systematic effects are minimized. Selecting muons with $p_T \leq 10$ GeV/c is likely to produce discrepancies between MC and data.

13.2 Muon Identification

Once a muon is required to have matching tracks in the inner and outer detectors, the main source of background consists of charged hadrons that leave a signature in the inner silicon tracker while also penetrating through the hadronic calorimeter and creating hits in the muon chambers. However, unlike muons, charged hadrons that penetrate the hadronic calorimeter and leave hits in the muon system will deposit significant energy in the calorimeters. Therefore, muon identification is based on an inside out algorithm that quantifies how well a muon track or global fit is compatible with that expected of a real muon. Calorimeter compatibility is quantified by taking

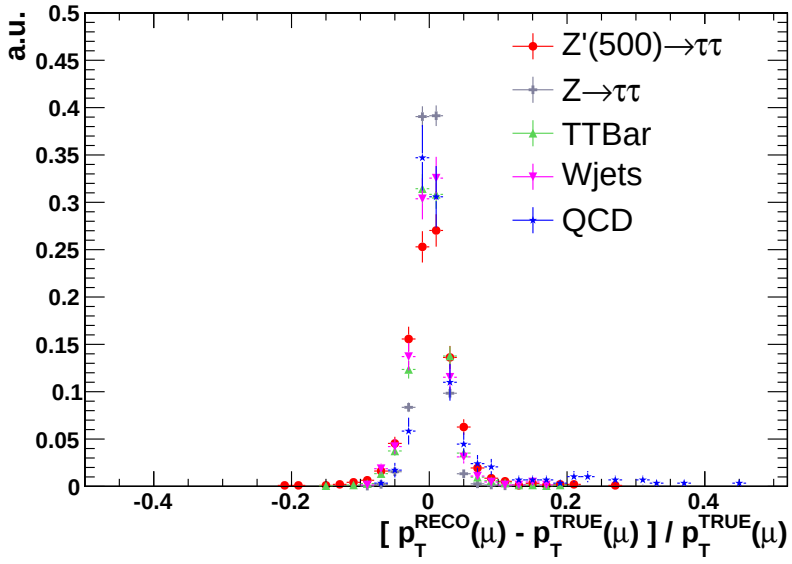

Fig. 13.1. Relative μ p_T resolution ($\frac{p_T^{RECO} - p_T^{TRUE}}{p_T^{TRUE}}$).

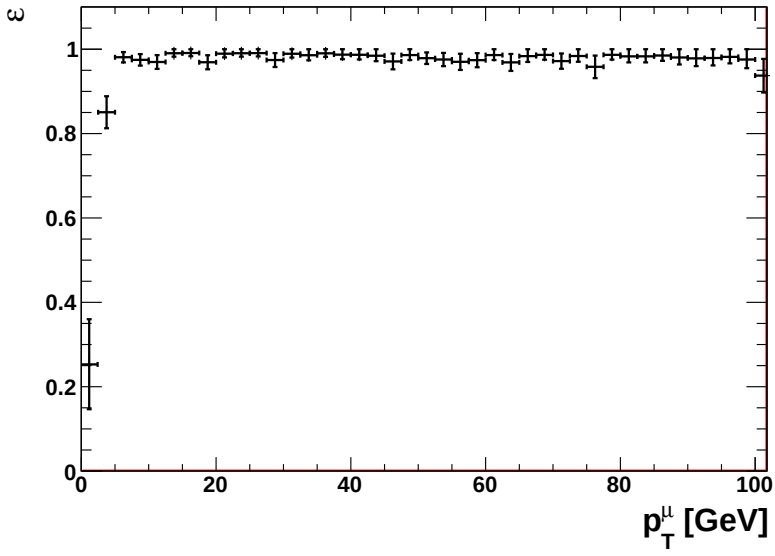

Fig. 13.2. Global muon reconstruction efficiency vs. p_T.

the muon trajectory and searching for energy deposits compatible with a *minimum ionizing* particle in the calorimeter. A *minimum ionizing* particle (MIP) is a particle whose energy loss rate as it travels through matter is near the minimum rate as described by the Bethe-Bloch equation [10]:

$$\frac{dE}{dx} = -k\frac{Z}{A}\frac{\rho}{\beta^2}[ln\frac{2m_e c^2\beta^2 E_M}{I_0^2(1-\beta^2)} - 2\beta^2] \tag{13.1}$$

Similarly, segment compatibility is quantified by extrapolating the silicon track outward and searching for compatible muon hits and segments in the muon subdetector. Specifically, the muon trajectory is extrapolated to the most likely location within each calorimeter volume and muon segment. Any calorimeter energy depositions and hits in the muon segments that are matched to the most likely positions are used to quantify the compatibility with a real muon signature. The matching is carried using the distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ between the most probable extrapolated track positions in the calorimeter (chambers) and the energy depositions (hits) in the surrounding region. The matching ΔR values range between 0.02 and 0.03 in the Ecal and 0.08 and 0.13 in the more coarsely segmented Hcal, depending on the η. The energy thresholds for observing the minimum ionizing calorimeter signals are chosen such that relative small minimum ionizing signals can be observed while also remaining above the noise level of the electronic readout systems. Energy depositions considered must have $E > 300$ MeV in Ecal and $E > 700$ MeV in Hcal. Once the associated depositions and segments are found, the compatibility value is calculated based on a likelihood function. For example, the calorimeter compatibility is calculated using a three-dimensional likelihood function of the form

$$\frac{P_S(x)\cdot P_S(y)\cdot P_S(z)}{P_S(x)\cdot P_S(y)\cdot P_S(z) + P_B(x)\cdot P_B(y)\cdot P_B(z)} \tag{13.2}$$

where P_S and P_B are the signal and background probabilities as functions of the measured energies in the Ecal (x), Hcal (y), and HO (z). The signal and background probability distributions are obtained from simulated samples of single-muons and pions, respectively [26–28]. Independent distributions are obtained for the barrel, overlap, and endcap regions of the detector and for different track p_T ranges. Because there is no HO scintillator layer in the endcap region of the detector, the $P_S(z)$ and $P_B(z)$ functions for this region are set to one. The resulting compatibility values for muons with $p_T > 20$ GeV/c from signal and various backgrounds relevant to this analysis is shown in Figures 13.3 and 13.4. Because the backgrounds relevant to this analysis are mostly due to processes where a muons are produced, Figures 13.3 and 13.4 does not show large discriminating power with the use of this variable. However, because the contributions from e.g. QCD is not expected to agree with the expectations from MC, the compatibility selections are maintained within the analysis to ensure robustness. Although the background contributions from pions that fake muons is expected to be small, maintaining the compatibility requirements ensure that pion based backgrounds such as QCD do not fluctuate upward by some large factor. In general, the cut values placed on calorimeter-compatibility and segment-compatibility are analysis dependent. In this analysis, the pion veto requirement is defined as a linear combination of the segment and calorimeter compatibilities:

$$\pi \text{ Veto} = 0.8 * C + 1.2 * S > 1 \tag{13.3}$$

where C and S are the segment and calorimeter compatibilities respectively.

Non-prompt muons can be produced from heavy flavour decays and decays in flights. However, unlike muons from processes such as $Z \rightarrow \tau\tau$, these particular muons are expected to be within jets. As discussed in previous sections, because hadronic jets tend to have wider energy profiles, isolation becomes an important

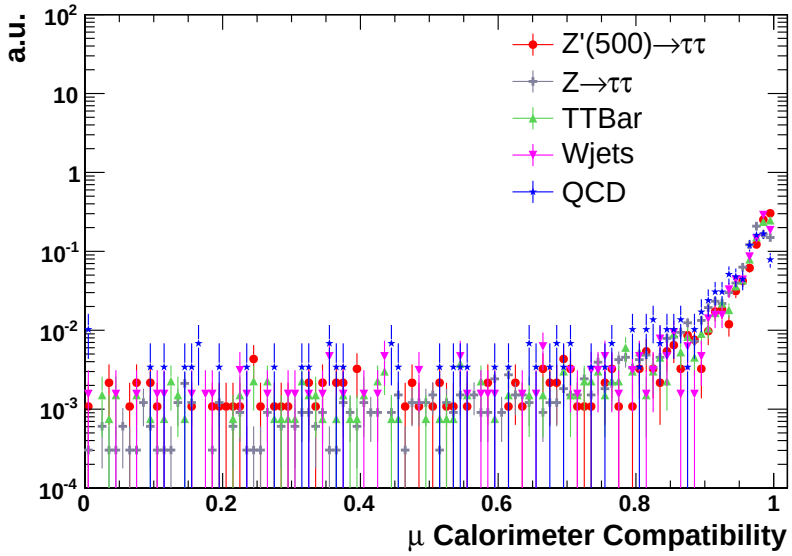

Fig. 13.3. Muon calorimeter compatibility for signal and various backgrounds.

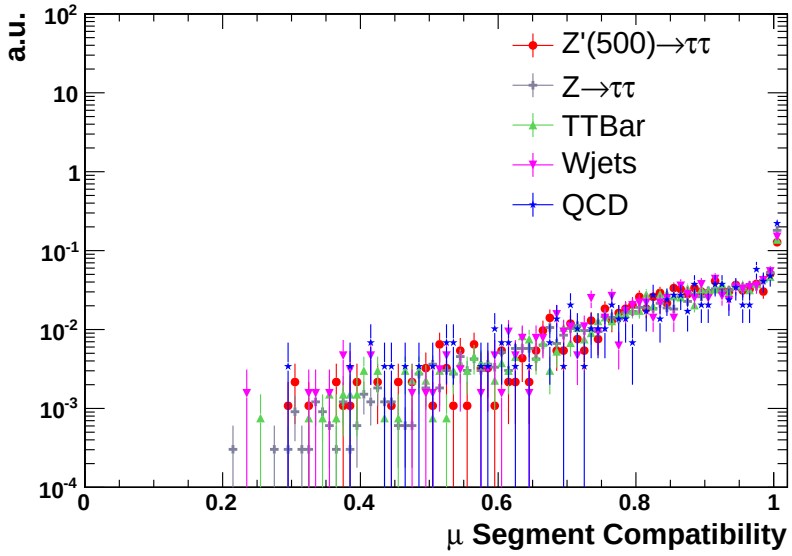

Fig. 13.4. Muon segment compatibility for signal and various backgrounds.

discriminator against backgrounds such as QCD. For muons produced by leptonic tau decays, tracks and calorimeter depositions around the muon trajectory are mostly a result of underlying events or pile-up particles. Therefore, an isolation requirement - requiring minimal calorimeter deposits and tracks within a region around the muon - is exploited as the main discriminator against muons from heavy flavour decays and decays in flight. Isolated muons are required to have minimal energy from tracks or Ecal depositions in a region of $\Delta R = 0.4$ around the lepton trajectory. Tracks considered in the calculation of isolation are required to be near the primary vertex. Isolation for muons is defined as:

$$I = \sum_i P_T^i \tag{13.4}$$

where the index i runs over all tracks (track isolation) or Ecal depositions (Ecal isolation). At CMS, the standard practice is to use isolation variables that are weighted by the muon p_T, $I_{rel} = I/p_T$. The main reason for using relative isolation is because it provides better discrimination against jets at low p_T ($p_T \sim 5 - 15$ GeV/c). However, this analysis is optimized to (1) be sensitive to high p_T muons from new hypothetical massive particles, and (2) minimize systematic effects. Therefore, because the muon p_T threshold ($p_T > 20$ GeV/c) is chosen to stay away from the *turn-on curve* of the reconstruction efficiency and because muons from heavy mass resonances are expected to have $p_T \sim 100$ GeV/c, *absolute* isolation is used instead. Absolute isolation has the additional advantage that it does not deal with systematic effects due to p_T and it provides muon identification efficiencies that do not depend on p_T. This can be seen in Figure 13.5, which shows the muon identification efficiencies as a function of muon p_T. Table 13.1 shows the complete list of μ identification criteria, while Figures 13.6 and 13.7 show the muon track and Ecal MC based isolation distributions for $Z' \to \tau\tau$ and backgrounds relevant to this analysis. One can see

that for $Z' \to \tau\tau$, most events reside in the lowest bin. Because $Z' \to \tau\tau$, $Z \to \tau\tau$, $Z \to \mu\mu$, and W+jets all have real clean muons, isolation distributions are similar. For the case of $t\bar{t}$, there are two contributions: (1) muons from the decay of W bosons, and (2) muons from the decay of b quarks. For this reason, $t\bar{t}$ has a similar structure as e.g. $Z \to \tau\tau$ and $Z \to \mu\mu$ in the low end of the distribution. However, a secondary jet structure is observed in the high end of the distribution. In the case of QCD, the only contribution is due to the heavy flavour decays such as the leptonic decays of $b\bar{b}$ pairs. Therefore, the isolation distribution for QCD sits mostly at high values of isolation. The probabilities for jets from a QCD muon enriched sample to fake a muon are shown in Figure 13.8. Because the QCD muon enriched samples are mostly $b\bar{b}$ events, the fake rates in Figure 13.8 represent the fake rates for heavy flavour decays and decays in flight. The fake rates are $< 10^{-2}$ for the entire range of p_T.

Table 13.1
μ Identification.

Cut
"Global" μ
Tracker hits ≥ 10
Pixel hits ≥ 1
≥ 2 chambers with matching segments
Global fit $\chi^2/NDOF < 10$
≥ 1 hit in muon system
$\lvert d_0 \rvert < 0.2$ cm
$0.8{*}C + 1.2{*}S > 1$
$\sum p_T$ of iso. tracks/Ecal RecHits < X (analysis dependent)

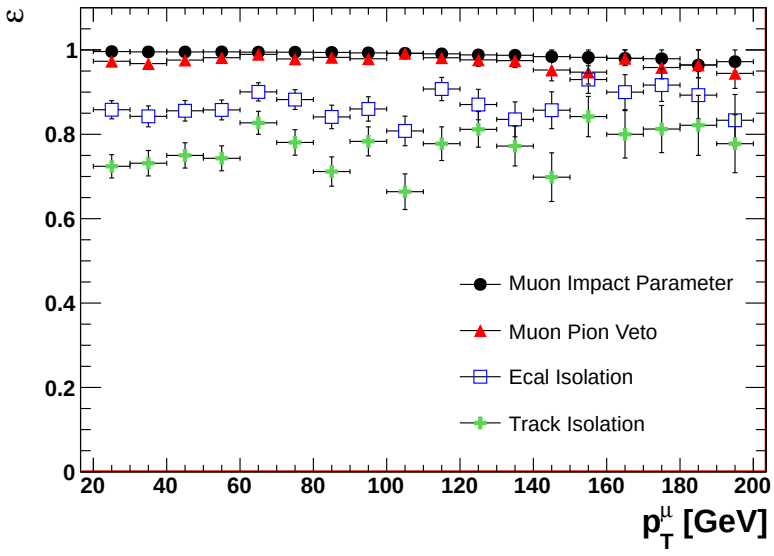

Fig. 13.5. Muon identification efficiency vs. p_T.

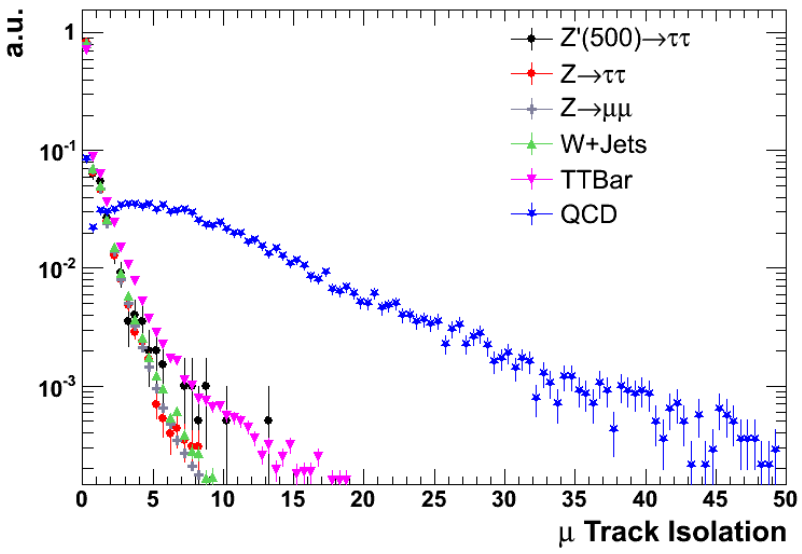

Fig. 13.6. $I = \sum_i P_T^i$ of tracks in the muon isolation region for signal and various backgrounds.

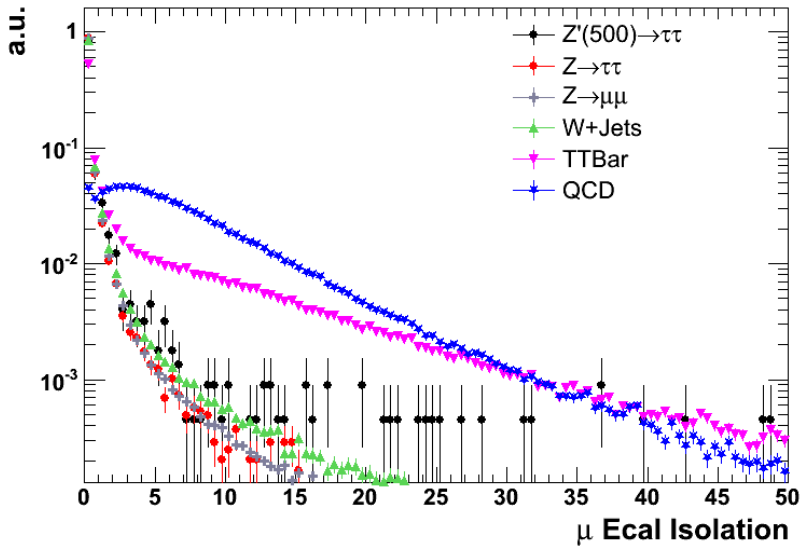

Fig. 13.7. $I = \sum_i P_T^i$ of reconstructed Ecal hits in the muon isolation region for signal and various backgrounds.

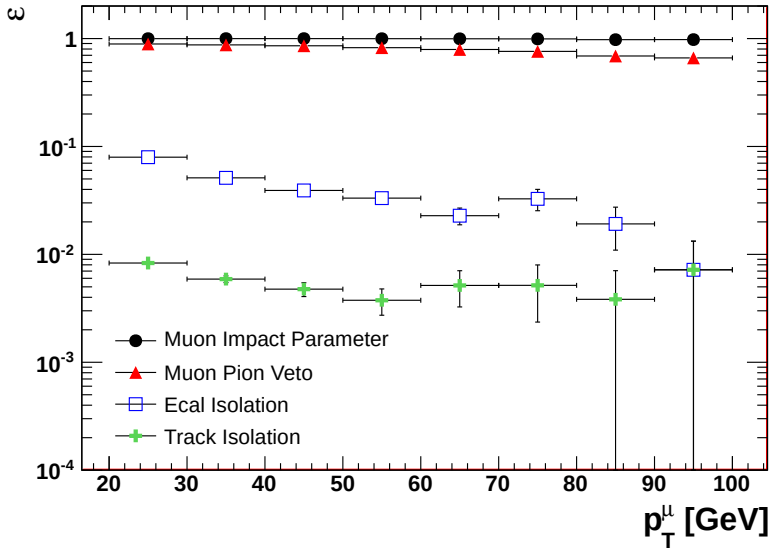

Fig. 13.8. Muon fake rates from heavy flavor decays and decays in flight vs. p_T obtained using a QCD muon enriched sample.

13.3 Particle Flow Reconstruction

Because τ's decay immediately after being produced, the existence of τ_h's must be inferred from the corresponding signatures in the CMS detector. There are many ways to reconstruct taus, one of which uses only calorimeter based energy deposits and tracking information to attempt to reconstruct a potential τ_h candidate. More recently, particle flow (PF) reconstruction techniques have been used to construct a mutually exclusive collection of reconstructed particles [29] - namely muons, electrons, photons, charged hadrons, and neutral hadrons - which are then used as input for the tau reconstruction algorithms.

Particle Flow reconstruction and identification is performed by using a combination of the information from each CMS subdetector, under the form of charged-particle tracks, calorimeter clusters, and muon tracks. Because charged particle momentum and position are measured with much better resolution in the tracker than the calorimeters, even for particle $p_T \sim 100$ GeV, the tracker is the centerpiece of the particle flow techniques.

An iterative tracking technique is used to ensure high efficiency of track finding and negligible fake rates at the desired p_T ($\sim$ a few hundred MeV). This is accomplished by seeding and reconstructing tracks using stringent criteria so as to achieve a negligible fake rate. This first iteration, however, does not achieve significantly high enough track finding efficiency. Therefore, during subsequent iterations, hits that are unambiguously assigned to the tracks are removed, and less stringent seeding criteria is used. This increases the probability to find a real track, while hardly increasing the probability to reconstruct a track that is not there (fake track). Only a few iterations are required to achieve a track finding efficiency of $\sim$ 99.5% for muons, and $> 90\%$ for charged hadrons from jets. Finally, in the last few iterations, tracks found in the first stages of the iterative tracking algorithm are required to

satisfy relaxed constraints on the origin of the vertex to allow the reconstruction of secondary charged particles from photon conversions, nuclear interactions in the tracker material, and the decay of long lived particles such as K_S^0's or Λ's [29].

Once tracks are reconstructed, the PF algorithm uses a calorimeter clustering algorithm [29] to (1) detect and measure the energy and direction of stable neutral particles such as photons and neutral hadrons; (2) separate energy deposits due to neutral particles from from charged hadrons; (3) reconstruct and identify electrons and all accompanying Bremsstrahlung photons; (4) improve the energy measurement of charged hadrons for which the track parameters were not determined accurately, which more often occurs in cases where the tracks have low quality fit parameters, or high p_T tracks. The calorimeter clustering algorithm is performed separately for Ecal Barrel (EB), Ecal Endcap (EC), Hcal Barrel (HE), and Hcal Endcap (HC). The clustering algorithm is not applied to the forward calorimeter (HF). First, any calorimeter cell with energy above some given energy threshold is identified as a "cluster seed." Next, neighboring cells with energy values that sit two standard deviations above the noise level, 100 MeV in the barrel and 300 MeV in the endcaps, are combined with the "cluster seeds" to form "topological clusters."

Because charged particles such as electrons will produce both tracks and calorimeter clusters, a linking algorithm must be employed to resolve any ambiguities and remove the double counting from different subdetectors. To establish a link, a track is extrapolated from its hits in the outer layers of the tracker to (1) the Ecal depth expected from a typical electron shower profile; (2) the Hcal depth expected from a hadron shower profile. A link is established if the extrapolated position is within the cluster boundary. In order to collect Bremsstrahlung energy from electrons, tangents to the tracks at each tracker layer are extrapolated to the Ecal with the requirement that the extrapolated tangents lie within the cluster boundary. Establishing a link

between the Ecal and Hcal clusters is achieved by requiring Ecal clusters to be within the Hcal cluster boundaries. Finally, in order to create a link between a track in the inner tracker layers and a track in the muon chambers, a global fit between the two tracks is performed and a link is determined if the χ^2 value is within the acceptable range (see section 13.1).

The final stage of the particle flow reconstruction is to construct the mutually exclusive collection of particles from the links between clusters, tracker tracks, and muon tracks. First, a "particle flow muon" is constructed if the global fit of the tracker and muon tracks results in a combined momentum that is within three standard deviations from the momentum of the track from the inner tracker. If a track passes the criteria for a particle flow muon, the link is removed from consideration. Remaining tracks considered for the reconstruction of charged hadrons are required to have relative p_T uncertainty that is within the calorimetric energy resolution expected for charged hadrons. Neutral hadrons are constructed by comparing the momentum of tracks with the energy detected in the clusters linked to tracks. If there is more energy in the linked cluster than is accounted for from the linked tracks (considering the resolution of the tracks and calorimeters), neutral hadron candidates and photon candidates are formed. The characteristic detector signatures of particles at CMS and the general strategy behind the particle flow reconstruction methods is depicted in Figure 13.9.

13.4 Tau Reconstruction

Although there are several algorithms that can be used for the identification and reconstruction of hadronically decaying tau letons [30], this thesis makes use of the simpler and more robust cut based approach suitable for early data. Tau reconstruction and identification algorithms use particle flow reconstruction to build hadronic tau jets. As discussed in the previous section, the PF algorithm builds a

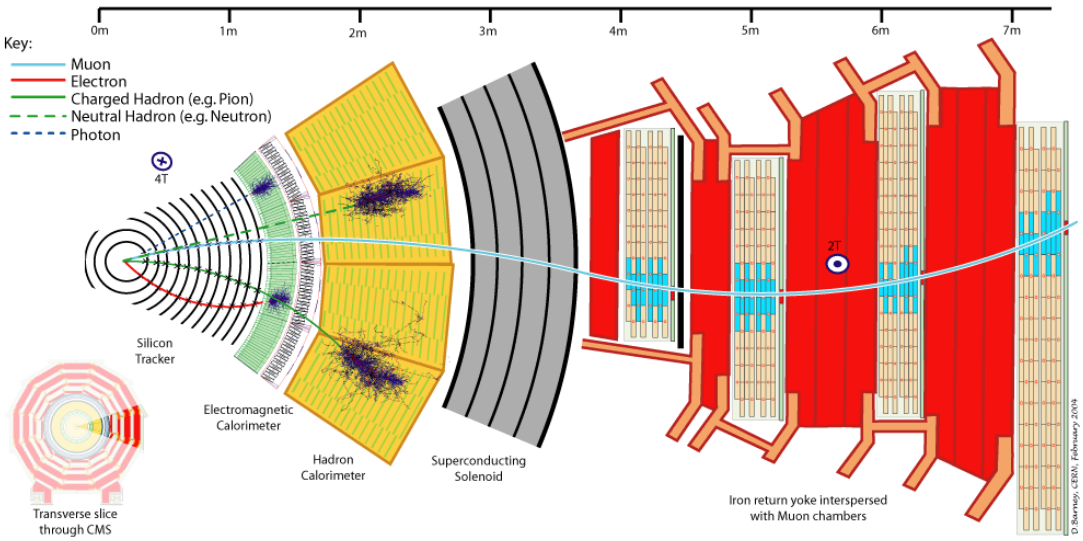

Fig. 13.9. Transverse slice of the CMS detector.

mutually exclusive collection of particles by combining tracks from the inner silicon tracker and muon system with calorimeter clusters to form PF candidates. Particle flow candidates are then used as inputs to build PF jets using the anti-k_t algorithm [31]. The anti-k_t algorithm is based on the calculation of two distances

$$d_{ij} = \min(\frac{1}{k_{t,i}^2}, \frac{1}{k_{t,j}^2})\frac{\Delta R_{ij}^2}{R^2} \tag{13.5}$$

$$d_{iB} = \frac{1}{k_{t,i}^2} \tag{13.6}$$

where $k_{t,i}$ is the transverse momentum of cluster i, R the jet reach parameter and ΔR_{ij} the distance in η–ϕ space between cluster i and cluster j

$$\Delta R = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2} \tag{13.7}$$

For the first iteration of the anti-k_t algorithm, d_{ij} is calculated for each pair ij, while d_{iB} is calculated for each cluster. In subsequent iterations, cluster pairs with the smallest value of d_{ij} are combined to form a "protojet" composed of the two clusters i and j and the clusters i and j are removed from the cluster for the next iteration.

The PF anti-k_t jet four-momentum is obtained by adding the four-momenta of all the associated jet PF candidates. Because tau jets tend to be more collimated with low particle multiplicity, PF τ's are built from PF jets by defining a narrow region, *signal cone*, around the central jet axis to define the tau constituents and a larger region to calculate variables such as isolation designed to discriminate τ_h's

from hadronic jets. Although the size of the signal and isolation regions can be analysis dependent, the "optimal" sizes should be determined as a compromise between the probability to correctly identify a τ_h and the probability to incorrectly identify a hadronic jet as a τ_h. The signal cone should be large enough to contain all the tau decay products and also be as small as possible to reject a large fraction of the QCD jet background.

To determine the optimal cone definition, it first needs to be established whether using a cone size of ΔR is a good choice for the general reconstruction of τ_h's. First consider a particle under a boost $\beta = v/c$ along the z-axis. The lorentz transformations of relativity tell us that the 4-momentum of (p_x, p_y, p_z, E) transforms in the following way:

$$q = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \gamma & \beta\gamma \\ 0 & 0 & \beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} \mathrm{p}_x \\ \mathrm{p}_y \\ \mathrm{p}_z \\ \mathrm{E} \end{pmatrix} = \begin{pmatrix} \mathrm{p}_x \\ \mathrm{p}_y \\ \gamma(p_z + \beta E) \\ \gamma(\beta p_z + E) \end{pmatrix}$$

The p_z and E components are the only components that transform under the lorentz boost. Therefore, the rapidity of the particle can be defined as follows:

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} \tag{13.8}$$

The rapidity transforms as $y \to y + \tanh^{-1}\beta$ under the boost in the z direction. Furthermore, because particles at the LHC are traveling at velocities close to the

speed of light, then $p >> m$ and $p_z/E \sim p_z/p = \cos\theta$. Using the trigonometric identity $\cos\theta = (1 - \tan^2\frac{\theta}{2})/(1 + \tan^2\frac{\theta}{2})$, the rapidity y can be shown to be fairly well approximated by the pseudorapidity $\eta = -\ln\tan\frac{\theta}{2}$. Therefore, the following transformations follow:

$$\phi \to \phi \tag{13.9}$$

$$\eta \to \eta + \tanh^{-1}\beta \tag{13.10}$$

This means that the distance in η-ϕ between particles within a jet are not changed under a boost along the z axis:

$$\Delta\phi = \phi_1 - \phi_2 \to \phi_1 - \phi_2 = \Delta\phi \tag{13.11}$$

$$\Delta\eta = \eta_1 - \eta_2 \to (\eta_1 + \tanh^{-1}\beta_1) - (\eta_2 + \tanh^{-1}\beta_2) \sim \Delta\eta \tag{13.12}$$

The previous calculations are based on the invariance of the transverse momentum given some boost in the z direction. Therefore, the following calculations can be interpreted as follows: *"Given the E_T of a jet, the boost factor along the z direction plays a negligible affect on the determination of jet shape variables."*

This is a powerful conclusion because it tells us that any τ_h variables that are used to discriminate against QCD hadronic jets are more robust if calculated using a constant ΔR. However, the above statements assume that the E_T is already known. Although the anti-k_t algorithm provides a very good determination of hadronic jet E_T, it is not the case for τ_h due to the "pencil like" decay structure that leads to a low

multiplicity of particles with significant energies. To determine the proper metric for the signal cone definition used to calculate the τ_h E_T and select the τ_h constituents, one can consider an idealized picture of the decay of tau leptons as illustrated in Figure 13.10. Writing the momenta of decay products in the tau rest frame as

$$P_{\nu_\tau} = \frac{m_\tau}{2}(0, \sin\theta^*, \cos\theta^*, 1) \tag{13.13}$$

$$P_{\pi^-} = \frac{m_\tau}{2}(0, \text{-}\sin\theta^*, \text{-}\cos\theta^*, 1) \tag{13.14}$$

and performing the Lorentz boost from the τ rest frame to the laboratory frame yields

$$\alpha = \arccos\left(\frac{p_\tau^2 \sin^2\theta^* - m_\tau^2}{p_\tau^2 \sin^2\theta^* + m_\tau^2}\right) \tag{13.15}$$

where α is the angle between the decay products in the laboratory frame and θ^* is the decay angle in the rest frame of the tau lepton. The angle between decay products (as observed in the detector) as a function of visible tau energy is shown in Figure 13.11 obtained using a MC sample of $Z \to \tau\tau$ events. The angle between the decay products in the detector is seen to decrease with increasing tau momentum/energy. Therefore, an energy or E_T dependent signal cone is the proper metric to define the correct tau constituents. The dashed line in Figure 13.11 shows an example of a "shrinking" cone and how it can successfully contain the tau decay products. For completeness, a similar depiction is shown in Figure 13.12 for the case of hadronic jets. Unlike taus, the distance between the constituent particles is not correlated with the energy of the jet. Therefore, a shrinking signal cone has an additional advantage that it can be used to provide further discrimination against

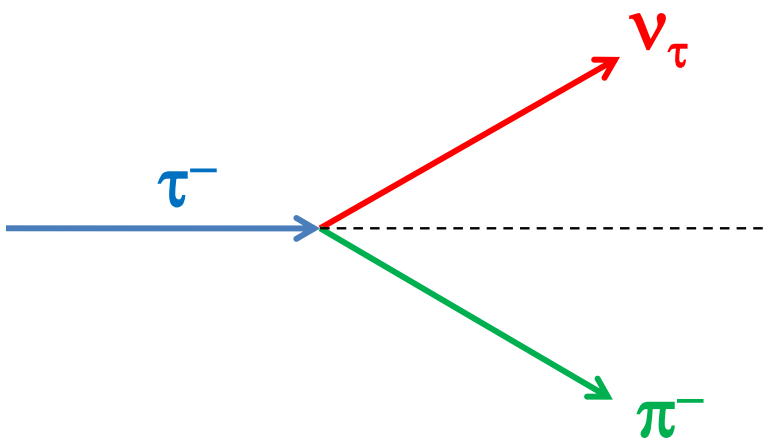

Fig. 13.10. Illustration of an idealized two-body decay of tau leptons.

hadronic jets when compared to larger fixed cones.

Based on the above considerations, τ_h's are built from PF jets by defining a narrow region (in $\eta-\phi$ space) with respect to the highest p_T track around the central jet axis (seed track). A E_T dependent signal region is used because it has the advantages presented above and also allows for the efficient reconstruction and identification of taus in the high mass region (high p_T^τ) as well as the low mass control region (low p_T^τ): $\Delta R_\tau^{Signal} = 5.0/E_T^{seed}$. It is important to note that although the most optimal signal cone can be defined by considering an E_T dependent cone using the geometrical angle α, a shrinking cone in η-ϕ space is used instead as it provides a consistent definition between the metric used to calculate jet shape quantities such as isolation ($\Delta R = 0.5$) and the metric used to define the tau constituents.

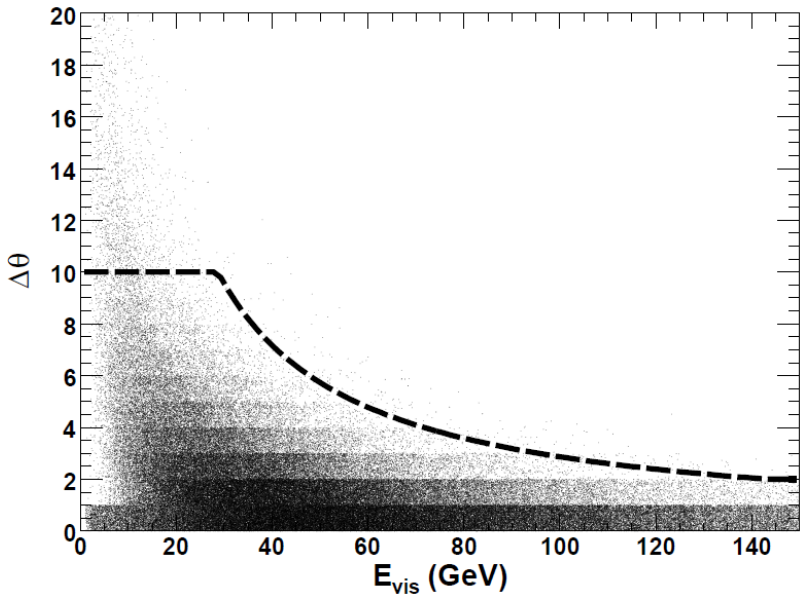

Fig. 13.11. $\Delta\phi$ vs. τ energy obtained using a simulated sample of $Z \to \tau\tau$ events.

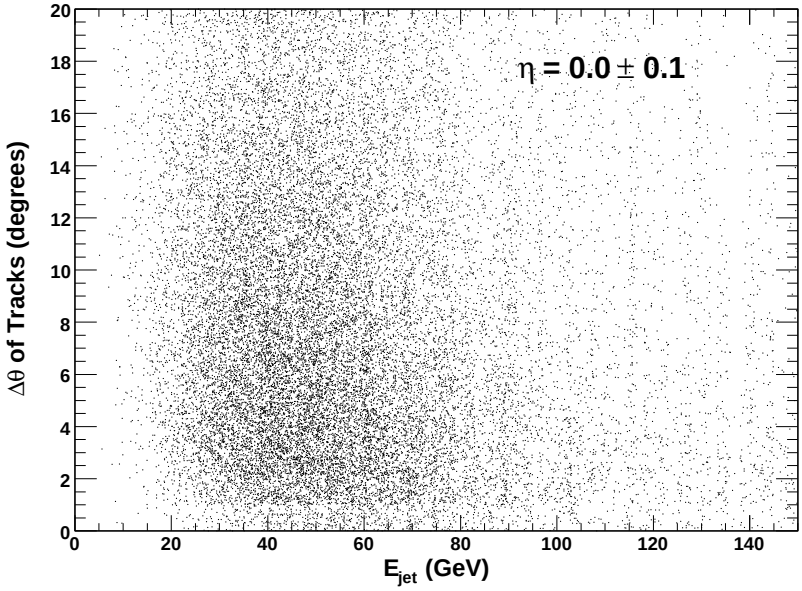

Fig. 13.12. $\Delta\phi$ vs. jet energy obtained using a simulated sample of $Z \to q\bar{q}$ events.

In the previous considerations, detector effects were not considered. Although a shrinking signal cone of $\Delta R_{\tau}^{Signal} = 5.0/E_T^{seed}$ for charged hadrons is still valid when one considers detector effects, the same cannot be said about neutral pions. Having a lifetime of 8.4 x 10^{-17} seconds, neutral pions can decay to a pair of photons almost immediately after being produced. As the photons traverse the tracker material, they can interact with the detector material and produce electron-positron pairs. Because the photon decayed within the tracker material, the electron-positron pairs will not be reconstructed as charged particles. Instead, the electron-positron pairs will deposit their energy in the Ecal, resulting in the reconstruction of PF photons. Furthermore, because these electrons bend in the electric field, they can be swept out of the signal region into the isolation annulus and spoil reconstruction and identification of taus. This effect is depicted in Figure 13.13. To account for this effect, we define separate signal regions for PF charged hadrons and PF photons:

- $\Delta R_{\tau}^{Tracker\ Signal} = 5.0/E_T^{seed}$
- $R_{\phi}^{Ecal\ Signal} = 0.15$
- $R_{\eta}^{Ecal\ Signal} = 0.07$
- $(\frac{\Delta\eta^{\gamma}}{R_{\eta}^{Ecal\ Signal}})^2 + (\frac{\Delta\phi^{\gamma}}{R_{\phi}^{Ecal\ Signal}})^2 < 1$
- $\Delta R_{\tau}^{Isolation} = 0.5$

PF candidates that reside in the signal regions are used to recalculate the four-momenta of the PF τ. Any associated particle flow candidates that fall outside the signal region and within the outer isolation region are labeled as PF tau isolation candidates. The sketch in Figure 13.14 depicts the τ_h signal and isolation cone definitions. Figure 13.15 displays the τ_h E_T resolution using PF based reconstruction with a shrinking signal cone $\Delta R_{\tau}^{Signal} = 5.0/E_T^{seed}$ and outer isolation cone $\Delta R = 0.5$. The resolution is compared to (1) calorimeter based reconstruction, and (2) calorimeter based reconstruction with hadronic jet based corrections. Particle

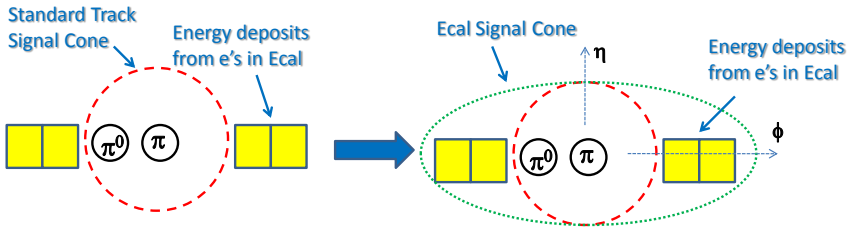

Fig. 13.13. Sketch depicting the effect of photon conversions on tau identification.

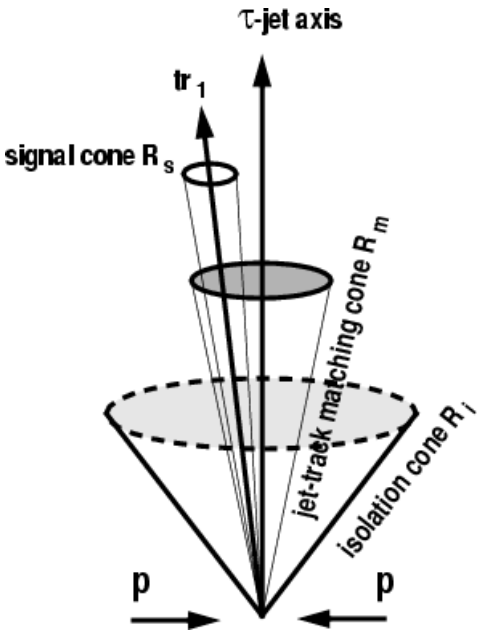

Fig. 13.14. Sketch depicting the τ signal and isolation cone definitions.

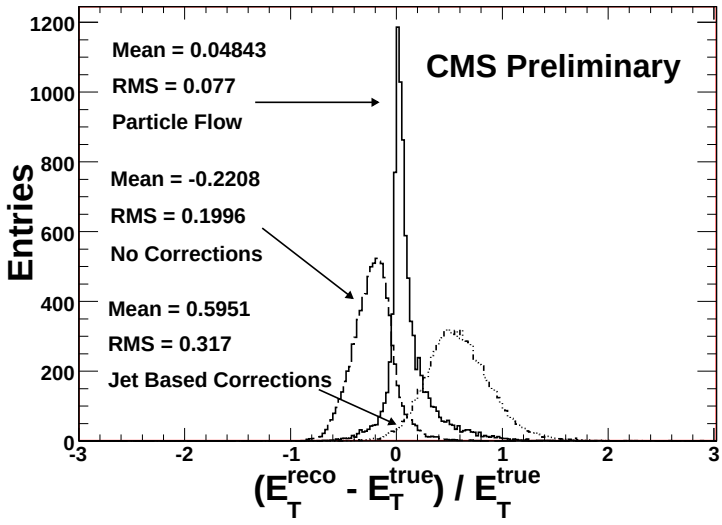

Fig. 13.15. Relative τ_h E_T resolution ($\frac{E_T^{RECO}-E_T^{TRUE}}{E_T^{TRUE}}$) for PF based reconstruction (solid line), calorimeter only based reconstruction, and calorimeter based reconstruction with jet based energy scale corrections.

flow based reconstruction provides significantly better resolution than calorimeter based reconstruction. Because τ_h's tend to be more collimated and contain smaller multiplicity of energetic particles, applying jet based corrections to the calorimeter based reconstruction of taus over-corrects the τ_h E_T and substantially worsens the resolution. The PF τ resolution is approximately 5%.

13.5 Tau Identification

The main difficulty in identifying hadronically decaying taus is devising a reliable discrimination technique against jets originating from the hadronization of quarks and/or gluons, namely QCD jets. The hadrons in a tau jet emerge from the decay of

a single particle ($W^{\pm}$) carrying large momentum, unlike the majority of QCD jets. Therefore, tau jets tend to be more collimated and characterized by low particle multiplicity. For this reason jet isolation, requiring no tracks or photons within a region centered in the jet direction, is exploited as the major discriminant against QCD jets.

The tau track isolation annulus is defined as the region between the track signal cone ($\Delta R = 5/E_T$) and the outer isolation cone ($\Delta R = 0.5$). Track isolation is calculated by adding the p_T of all particle flow charged hadrons above some p_T threshold in the isolation annulus:

$$I^{Track} = \sum_i p_T^i \tag{13.16}$$

where the index i runs over all PF charged hadrons. Similarly, the tau Ecal isolation annulus is defined as the region between the Ecal signal cone (ellipse with $R_\eta = 0.07$, $R_\phi = 0.15$) and the isolation cone. Ecal isolation is calculated by adding the p_T of all particle flow photons above some threshold in the Ecal isolation annulus:

$$I^{Ecal} = \sum_i p_T^i \tag{13.17}$$

where the index i runs over all PF photons. Choosing a threshold for isolation candidates is analysis dependent, but should be chosen as a middle ground between the level of inefficiency due to UE/PU and the tolerable probabilities for a jet to fake a tau. Because tau jets are more collimated and characterized by low particle multiplicity, taus are required to have minimal energy in the isolation annulus.

Jets that have minimal energy in the isolation annulus, and thus pass the isolation requirement can be discriminated against by considering the track multiplicity in the signal region. Hadronically decaying taus predominantly decay to one or three charged pions and neutral pions (see Table 7.1), whereas jets tend to have larger signal track multiplicities (see Figure 8.4). Therefore, taus are generally required to have one or three particle flow charged hadrons in the signal region. Table 13.2 shows the minimal set of quality cuts used to define signal and isolation PF charged/gamma candidates, while Figures 13.16 and 13.17 show the tau Ecal and track MC based isolation distributions for $Z' \to \tau\tau$ and backgrounds relevant to the this analysis. Figure 13.18 shows the MC based distributions for the number of tracks in the tau signal region for $Z' \to \tau\tau$ and backgrounds relevant to this analysis. At CMS, the standard practice for τ_h related analyses is to require τ_h candidates to have one or three charged hadrons in the signal region. However, in this analysis, τ_h candidates are required to have exactly one charge hadron in the signal region. The motivation behind using *one prong* taus is to: (1) minimize systematic effects from three prong taus, which are much more difficult to understand and measure; (2) significantly reduce the background contamination. The systematic effects for taus will be discussed in section 20.1. The discriminating power for the one prong choice of taus can be determined by considering Figure 13.18. For the case of signal $Z' \to \tau\tau$, the ratio between one prong taus n_1^τ and three prong taus n_3^τ is approximately 0.15 / 0.75 = $\frac{1}{5}$. For the case of jet backgrounds such as W+jets, the ratio between one prong jets n_1^{jet} and three prong jets n_3^{jet} is approximately 0.15 / 0.30 = $\frac{1}{2}$. Therefore, requiring one prong taus (as opposed to three prong taus) improves the signal to background ratio by approximately

$$\frac{R_1}{R_{1||3}} = \frac{\frac{n_1^\tau}{n_1^{jet}}}{\frac{n_1^\tau + n_3^\tau}{n_1^{jet} + n_3^{jet}}} \sim \frac{5}{2} \tag{13.18}$$

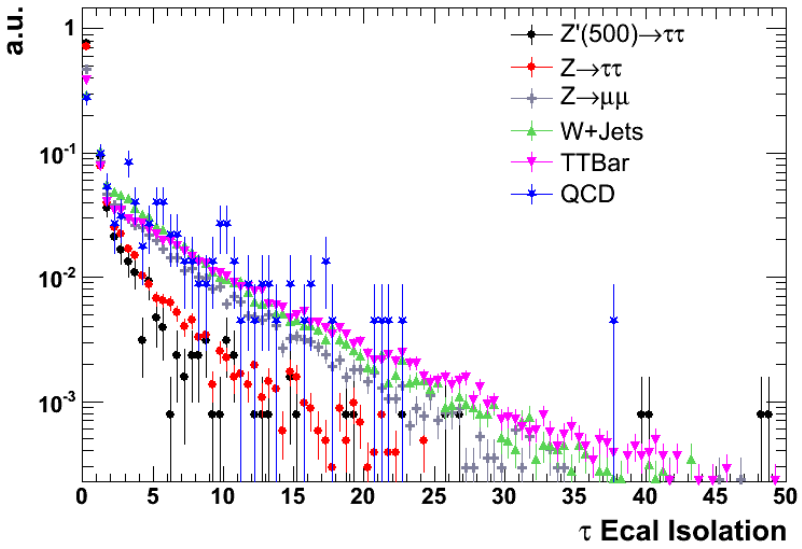

Fig. 13.16. $I = \sum_i P_T^i$ of PF photons in the tau isolation region for signal and various backgrounds.

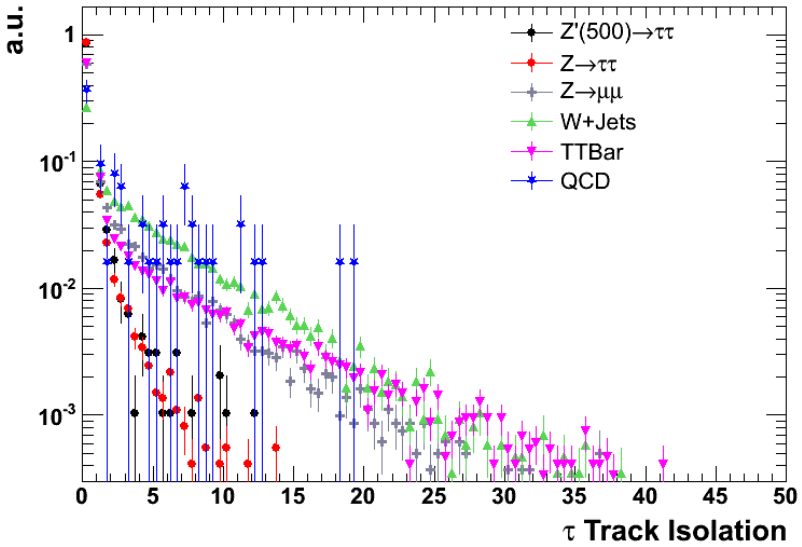

Fig. 13.17. $I = \sum_i P_T^i$ of PF charged hadrons in the tau isolation region for signal and various backgrounds.

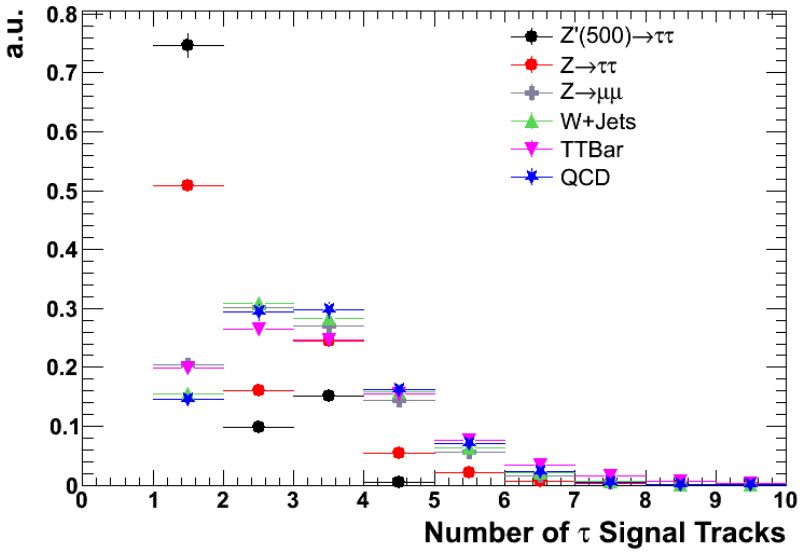

Fig. 13.18. Number of tracks in the τ signal region. For $Z' \to \tau\tau$, most taus contain one or three signal tracks.

Table 13.2
Quality cuts for τ signal and isolation candidates.

Quality Cut
Signal track $p_T > 0.5$ GeV
Isolation track $p_T > 1.0$ GeV
Track $\chi^2 < 100$
Track $\lvert d_{xy}\rvert < 0.03$ cm
Track $\lvert\Delta z\rvert < 0.2$ cm
Signal track hits > 3
Isolation track hits > 8
Signal γ $E_T > 0.5$ GeV
Isolation γ $E_T > 1.0$ GeV

Because inefficiencies in the isolation of tau jets are largely due to underlying events and/or pileup, as is the case for electrons and muons (see section 15 and Figures 15.3 and 15.4), the above requirements do not discriminate tau jets from electrons and muons. To reject muons, the tau seed track is required to have compatible hits or energy deposits with that expected of a pion:

$$\mu \text{ Veto} = 0.8 * C + 1.2 * S < 1 \tag{13.19}$$

Note that this is merely the inverted pion veto cut defined above for the case of muons. To discriminate against electrons, taus are required to have sufficient energy (with respect to the seed track momentum) arising from the hadronic calorimeter in a 3x3 region around the seed track (e.g. $H_{3x3}/P > 0.03$). Table 13.3 summarizes the entire set of τ identification criteria. The τ_h identification efficiencies as a function of τ_h p_T are shown in Figure 13.19. The tau tagging efficiencies are approximately 55% for signal $Z' \to \tau\tau$. The efficiencies are constant as a function of p_T, ensuring that systematic effects on tau identification due to uncertainties in tau energy scale

are minimized. The probabilities for jets from a QCD muon enriched sample to fake a τ_h are shown in Figure 13.20. Because the QCD muon enriched samples are mostly $b\bar{b}$ events, the fake rates in Figure 13.20 represent the fake rates for heavy flavour decays and decays in flight. The fake rates are $\sim 10^{-2}$ for the entire range of p_T. To get a more general idea of the fake rates for a broader class of jets, the fake rates are determined for hadronic jets from the simulated sample of W+jets (Figure 13.21). The overall jet$\rightarrow \tau_h$ fake rate using the sample of W+jets is approximately 10^{-3}–10^{-2}.

Table 13.3
τ Identification Criteria

Cut
τ seed track $p_T > 5$ GeV/c
τ seed track number of hits > 11 AND $H_{3x3}/P > 0.03$ (electron veto)
$0.8 * C + 1.2 * S < 1$ requirement on the τ seed track (μ veto)
1 charged hadron in the tracker signal region
$\sum P_T$ of isolation tracks/γ's $<$ X (analysis dependent)

13.6 Missing Transverse Energy

The typical energies of particles traveling outside the detector acceptance region are much less than the collision beam energies. Therefore, the CMS detectors allow for a fairly precise measurement of the momentum imbalance in the transverse direction relative to the beam direction (missing transverse energy, MET). Because the LHC is expected to probe theoretical extensions of the Standard Model (SM), large efforts have been placed on the development and improvement of variables that will be vital for the discovery of new physics. In particular, the measurement of a large transverse momentum imbalance at CMS could be strong evidence of new physics. In searches for new heavy mass resonances decaying to tau leptons, tau lepton energies are typically on the order of hundreds of GeV. Neutrinos from these

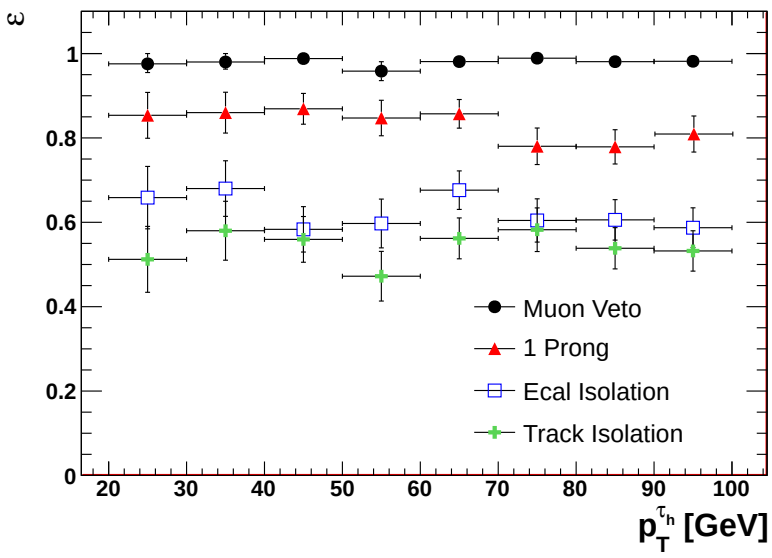

Fig. 13.19. τ_h identification efficiency vs. p_T.

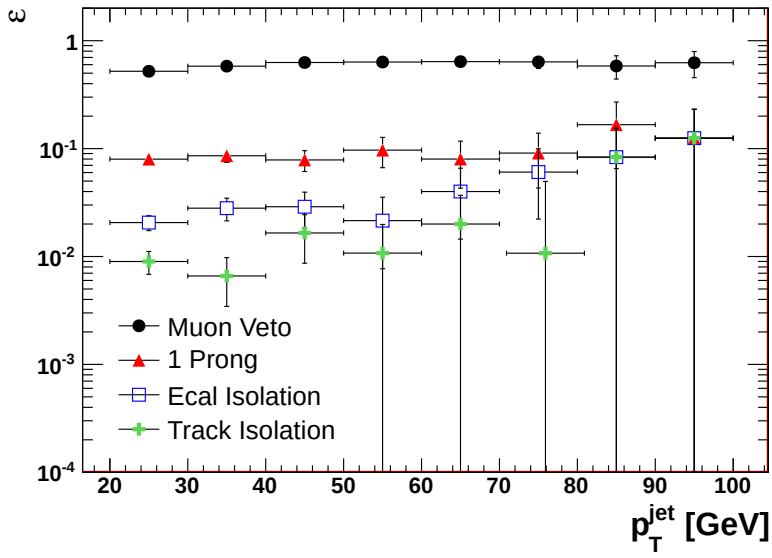

Fig. 13.20. τ_h fake rates from heavy flavor decays and decays in flight vs. p_T obtained using a QCD muon enriched sample.

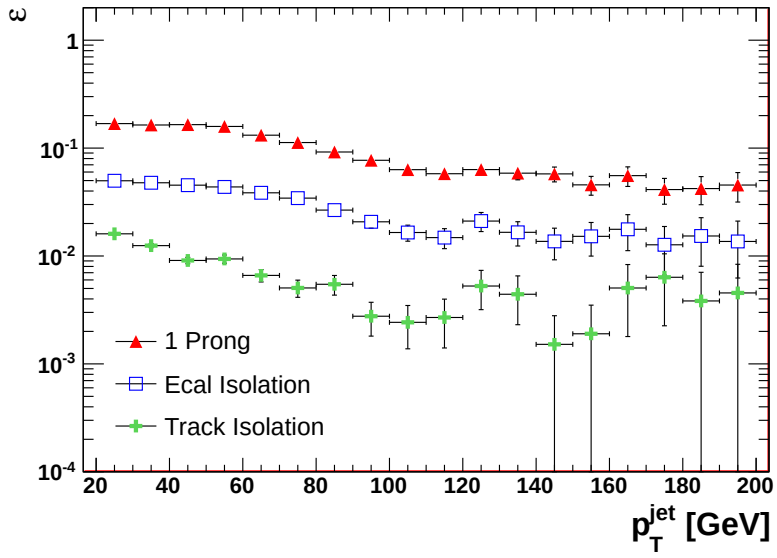

Fig. 13.21. jet$\rightarrow \tau_h$ fake rates vs. p_T using jets from a W+jets sample.

tau decays escape undetected and can leave an average momentum imbalance of 100 GeV (depends on τ_h p_T). Therefore, not only is the missing transverse energy an important variable for electroweak measurements and for searches for new physics, but also for the identification of final states involving tau leptons. As we shall see, when attempting to reconstruct a new mass resonance, any sign of new physics will usually reside in the tails of mass distributions. Since tau leptons decay to neutrinos, a good understanding of the tails of the momentum imbalance measurement is vital.

At CMS, the MET vector is mostly calculated by using energy depositions in the calorimeters (CaloMET) [32] or by using reconstructed particle momenta obtained by combining information from all the CMS subdetectors (PFMET) [33]. Because the initial momentum of the system at the LHC is $\overrightarrow{p} = \overrightarrow{0}$, the momentum imbalance in the detector can be calculated by determining the total visible momentum in the detector as follows:

$$\overrightarrow{p} = \overrightarrow{0} = \sum_i \overrightarrow{p}_i^{\text{visible}} + \sum_i \overrightarrow{p}_j^{\text{not visible}} \tag{13.20}$$

$$\sum_i \overrightarrow{p}_j^{\text{not visible}} = -\sum_i \overrightarrow{p}_i^{\text{visible}} \tag{13.21}$$

Therefore, CaloMET is defined as the transverse vector sum over raw uncorrected energy deposits in the calorimeter towers:

$$\overrightarrow{\not{E}_{\text{T}}} = -\sum_i (E_i sin\theta_i cos\phi_i \hat{\mathbf{i}} + E_i sin\theta_i sin\phi_i \hat{\mathbf{j}}) \tag{13.22}$$

where the index i runs over all calorimeter towers. Raw (uncorrected) CaloMET may be corrected to account for effects such as mis-measurements in jet energies,

and muons which deposit very little energy in the calorimeters. PFMET makes use of the more complex particle flow algorithms to reconstruct the momenta of individual particles:

$$\overrightarrow{\not{E}_{\mathrm{T}}} = -\sum_{i} \overrightarrow{p_i} \tag{13.23}$$

where the index i runs over all particle flow candidates. Figure 13.22 shows the $\not{E}_{\mathrm{T}}$ resolution for PF based reconstruction, calorimeter only based reconstruction, and calorimeter based reconstruction with jet based energy scale corrections obtained using the $Z \to \tau\tau$ MC simulated sample. Particle flow based $\not{E}_{\mathrm{T}}$ reconstruction provides significantly better resolution than calorimeter based reconstruction. The resolution is $\sim$ 5 GeV for particle flow based reconstruction and $\sim$ 10 GeV for calorimeter based reconstruction.

MET is a quantity sensitive to several detector effects such as noise from electronics, pile-up (PU), underlying-events (UE), cracks, and dead material. This means that disentangling and understanding the various systematic effects from individual sub-detectors can be very challenging. The analysis of cosmic ray data, beam splash data, and early collision data has allowed us to the test the performance of missing transverse energy. These studies have shown the major source of discrepancies between MC and data are due to anomalous behaviour from the hybrid photodiodes (HPDs), problems with the readout boxes (RBX), and large energy readouts from the electromagnetic crystals ("Ecal spikes"). ECAL spikes are characterized by very large energy readouts arising from a single crystal. Therefore, Ecal spikes are removed from the calculation of PFMET by applying a cut on S_4/S_1, which calculates the relative energy in 4 neighboring crystals around around the seed cell with respect to the energy of the seed cell. HPD discharge occurs when there is misalignment with

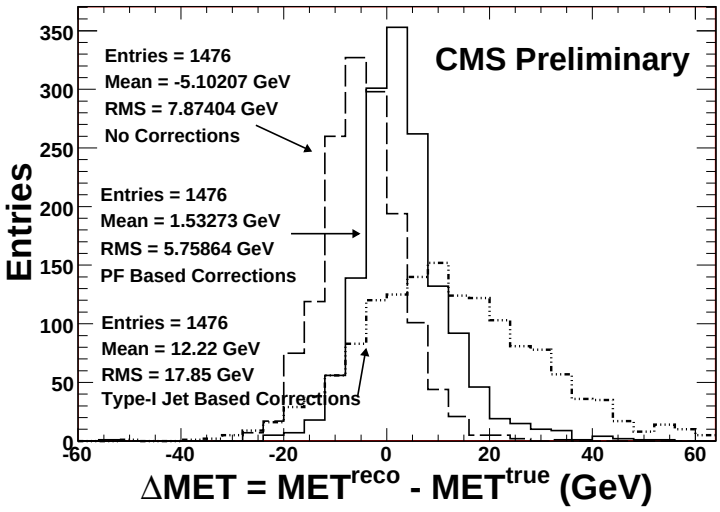

Fig. 13.22. MET resolution for PF based reconstruction, calorimeter only based reconstruction, and calorimeter based reconstruction with jet based energy scale corrections.

the magnetic field. This misalignment causes electrical discharge from the walls of the HPDs which results in large energy deposits in all 18 channels in the HPD. RBX noise is characterized by anomalous behavior in all four HPDs within a RBX. To remove events due to HB/HE noise, HPD hit multiplicity and timing variables are used to determine the compatibility with a real particle signature. The use of noise cleaning quantities in any analysis requires careful monitoring of the anomalous behavior. Because events due to detector anomalies such as Hcal noise mostly arise due to the imperfections of the detector, the discriminating power for noise rejection and cleaning variables can vary significantly as the detector conditions change. Additionally, imposing such requirements on an analysis can impose significant biases. Therefore, a monitoring scheme has been designed to select clean samples of anomalous events so that noise rates can be measured and the effectiveness of discriminating variables can be studied. The noise monitoring methodology is depicted in Figure 13.23. Figure 13.24 shows the HB/HE noise rate as a function of $\not\!\!E_{\mathrm{T}}$ threshold obtained using the monitoring methodology depicted in Figure 13.23. Similarly, Figure 13.25 shows the noise rate as a function of run number. The noise rates for $\not\!\!E_{\mathrm{T}} > 20$ GeV are determined to be $\sim 10^{-4}$–10^{-3} Hz for the entire data taking period. Therefore, even at large values of $\not\!\!E_{\mathrm{T}}$, the noise rate does not present a problem for this analysis. However, noise rejection variables are maintained within the analysis to ensure robustness. Figure 13.26 shows the RBX hit multiplicity, while Figure 13.27 shows S_4/S_1 as a function of Ecal seed cell energy. The three peak structure for the number of RBX hits obtained from a clean sample of anomalous events depicts the different types of Hcal anomalous behavior. The presence of "Ecal spikes" can be seen at large values of Ecal seed energy and $S_4/S_1 \sim 0$.

The simulated samples used for this analysis do not include the simulation of pile-up. Therefore, any $\not\!\!E_{\mathrm{T}}$ related variable will show discrepancies in data and MC. For example, Figure 13.28 displays the $\not\!\!E_{\mathrm{T}}$ distributions for data and MC in the $\mu\tau_h$

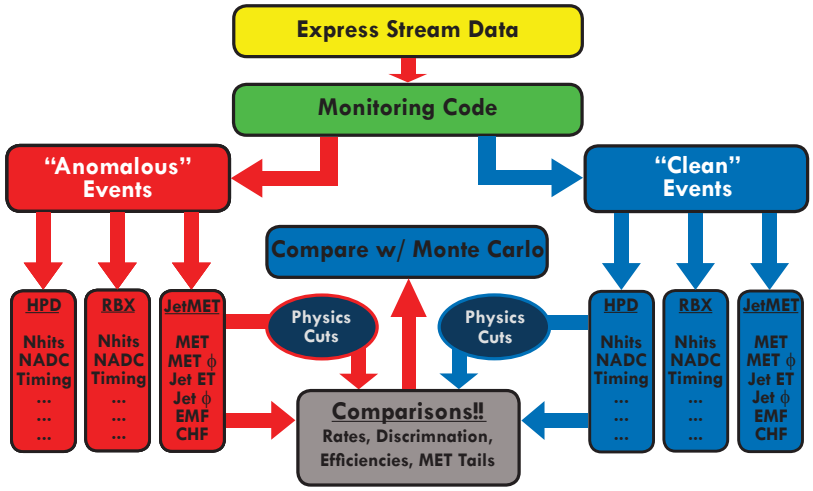

Fig. 13.23. Depiction of the $\not\!\!E_{\rm T}$ monitoring methodology.

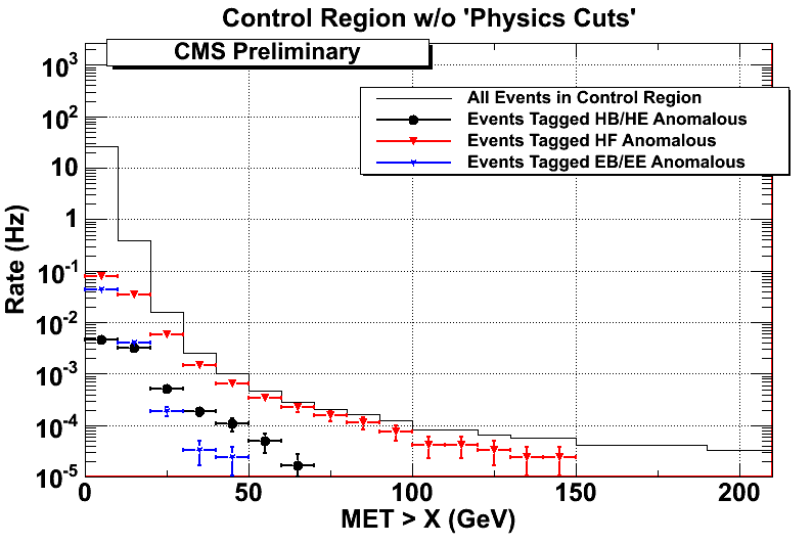

Fig. 13.24. Note Rate vs. $\not\!\!E_{\rm T}$ threshold.

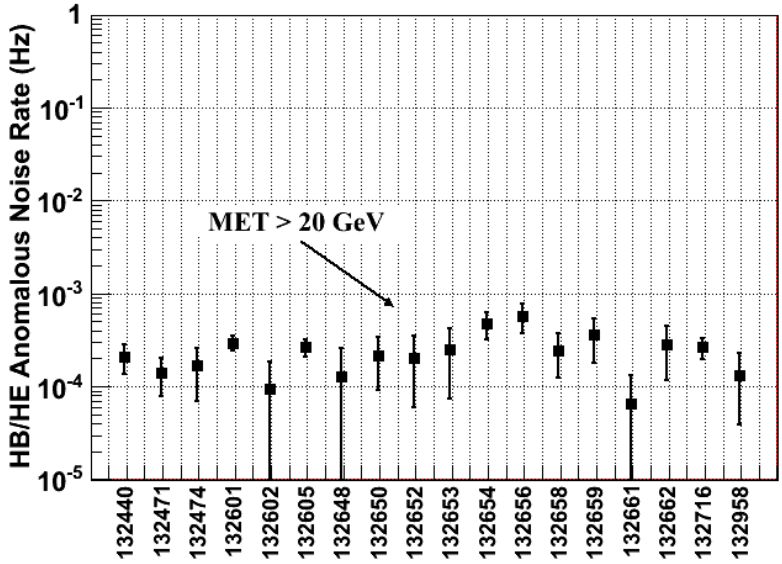

Fig. 13.25. Note Rate vs. Run. This figure represents the stability of the noise rates over the data taking period.

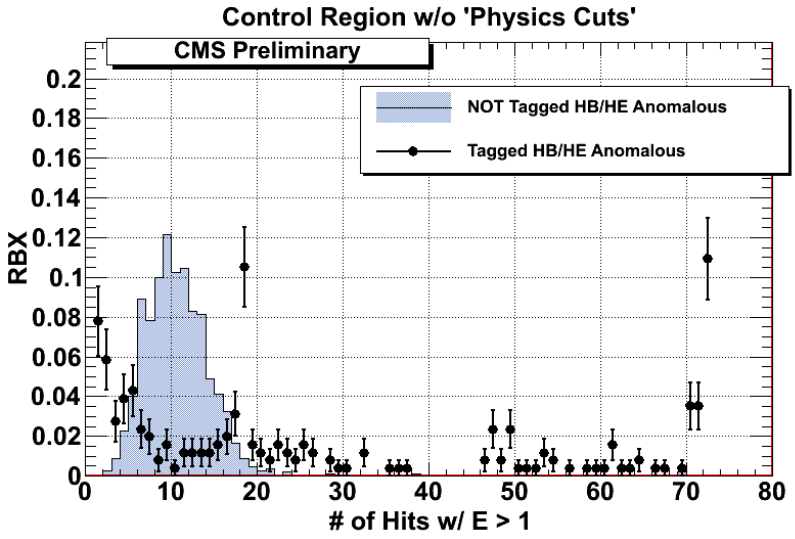

Fig. 13.26. Number of hits with $E > 1$ GeV within a Hcal readout box (RBX).

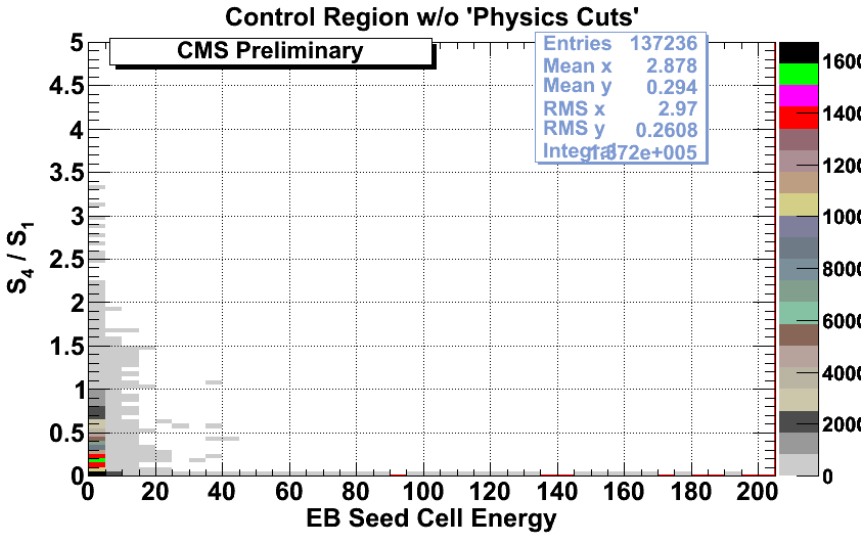

Fig. 13.27. S_4/S_1 - relative energy of a seed cell in Ecal with respect to the energy in neighboring cells.

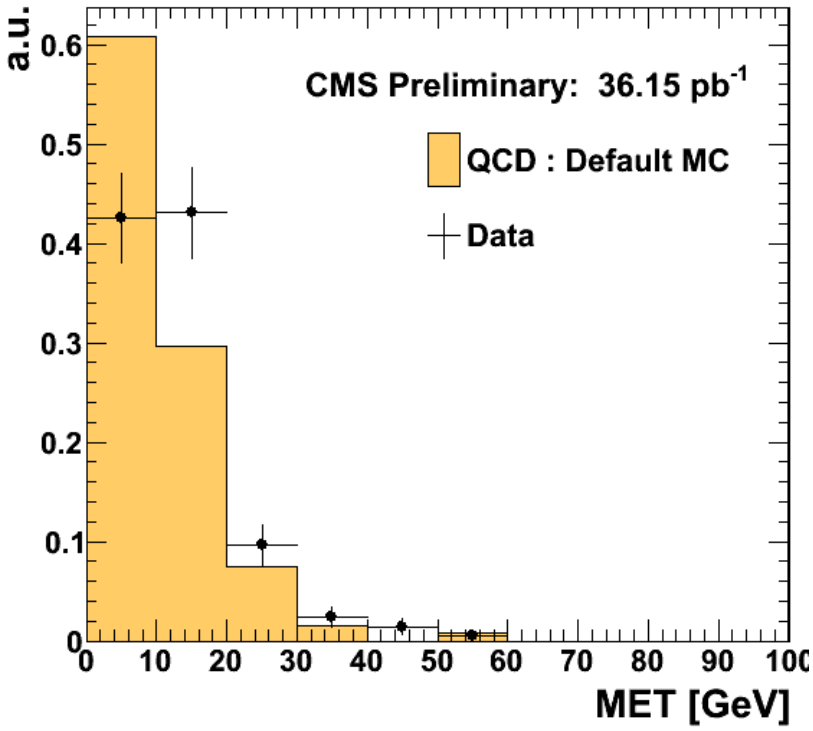

Fig. 13.28. $\not\!\!E_{\mathrm{T}}$ in $\mu\tau$ QCD control region without PU correction.

QCD control region (see section 17.1 for a definition of the QCD control region). One can clearly see a harder spectrum in data compated to MC. To account for pile-up in simulation, we recalculate the $\not\!\!E_{\mathrm{T}}$ as follows:

$$\overrightarrow{\not\!\!E_{\mathrm{T}}}^{Corrected} = \overrightarrow{\not\!\!E_{\mathrm{T}}}^{Raw} + \overrightarrow{\Delta \not\!\!E_{\mathrm{T}}}^{PU} \tag{13.24}$$

The pile-up corrections to $\not\!\!E_{\mathrm{T}}$ can be measured by obtaining a clean sample of γ+jet events. This topology has the advantage that there is no real intrinsic $\not\!\!E_{\mathrm{T}}$. Therefore, measurement of non-zero values of $\not\!\!E_{\mathrm{T}}$ are mostly due to (1) mis-

measurement of the jet energy/momentum, (2) mismeasurement of the photon energy/momentum, and (3) pile-up contributions. Therefore, the projection of $\not\!\!E_{\mathrm{T}}$ to the direction of the γ+jet pair quantifies the $\not\!\!E_{\mathrm{T}}$ resolution due to the mismeasurement of the γ and jet. Projecting the $\not\!\!E_{\mathrm{T}}$ in a direction perpendicular to the γ+jet pair quantifies the $\not\!\!E_{\mathrm{T}}$ resolution due to pile-up effects. The methodology used to measure the $\not\!\!E_{\mathrm{T}}$ resolution in this way is depicted in Figure 13.29. The $\not\!\!E_{\mathrm{T}}$ resolution was measured in [34] following this approach. The $\not\!\!E_{\mathrm{T}}$ resolution is parameterized in terms of the number of reconstructed primary vertices, n, per event.

$$\delta\not\!\!E_{\mathrm{T}} = \sqrt{n} \cdot \sigma_{PU} \cdot F_{scale}(\not\!\!E_{\mathrm{T}}) \tag{13.25}$$

In the equation above, σ_{PU} represents the PU resolution, while $F_{scale}(\not\!\!E_{\mathrm{T}})$ accounts for scale corrections. The pile-up resolution σ_{PU} in [34] is measured 3.53 $\pm$ 0.04 GeV. Figure 13.30 shows the number of reconstructed vertices per event in the QCD control region. To accurately apply the PU correction to the calculation of $\not\!\!E_{\mathrm{T}}$, the probability for an event to contain n primary vertices is extracted from each control region. Once the PU correction to $\not\!\!E_{\mathrm{T}}$ is included, data and MC distributions agree (Figure 13.31).

13.7 Di-Tau Mass Reconstruction

Because tau leptons decay to neutrinos which leave the detector undetected, one cannot fully reconstruct the mass resonance with the visible τ decay products. Additionally, because the invariant mass for background processes such as QCD are typically steeply falling distributions in the tails (where new mass resonances are expected), it becomes important to make use of $\not\!\!E_{\mathrm{T}}$ to attempt to separate signal from background and reconstruct the true mass resonance. Historically, several methods

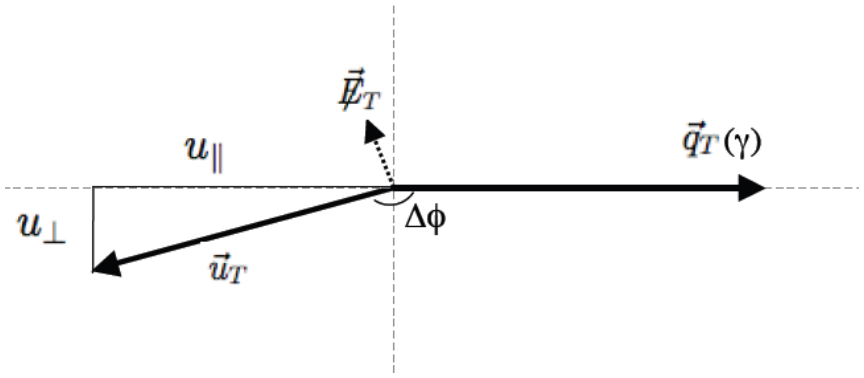

Fig. 13.29. Sketch depicting the the definition of the kinematic variables: photon momentum (q_T), recoil momentum (u_T) and it's projections ($u_\parallel, u_\perp$).

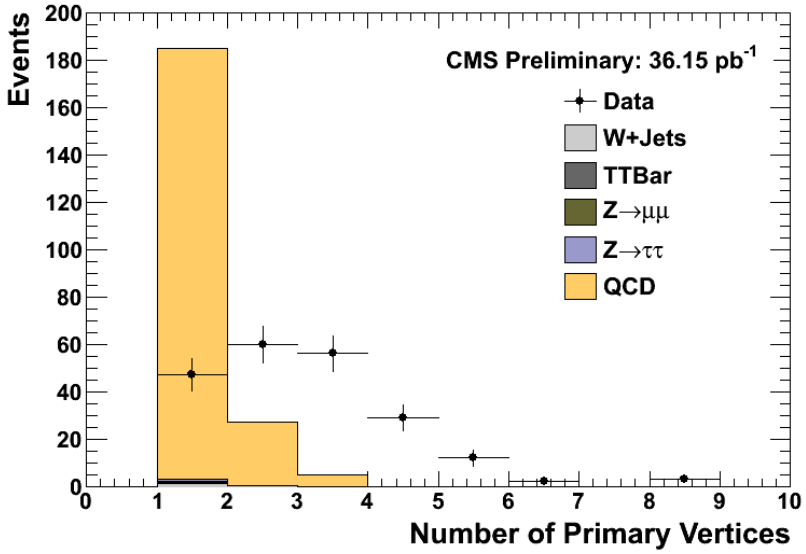

Fig. 13.30. Number of primary vertices in data and MC in the $\mu\tau$ QCD control region.

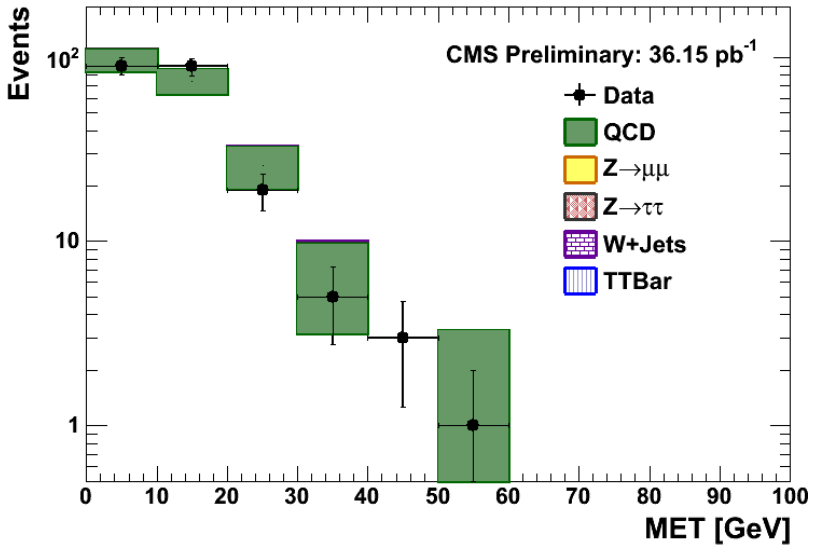

Fig. 13.31. $\not\!\!E_{\mathrm{T}}$ in the $\mu\tau$ QCD control region with the PU correction.

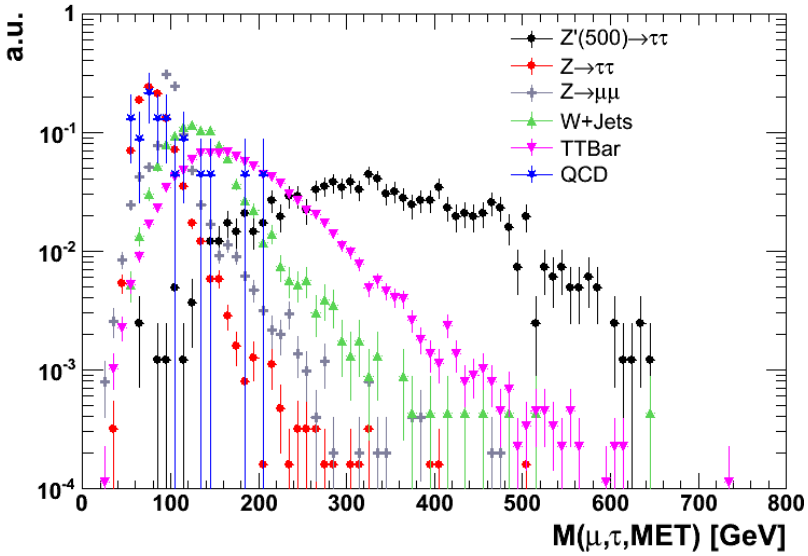

Fig. 13.32. $M(\mu, \tau, \not{E}_{\mathrm{T}})$ for signal Z' and various backgrounds.

such as the collinear approximation have been employed to reconstruct the true mass resonance. However, for the analysis presented in this thesis, one of the main sources of backgrounds is W+jet(s). In this case, the analysis achieves a high degree of sensitivity by requiring the tau candidates to be back-to-back in ϕ (see sections 13.8 and 16). This is precisely the regime in which the collinear approximation fails. Therefore, the mass is reconstructed as follows:

$$M(\tau_1, \tau_2, \not{E}_{\mathrm{T}}) = \sqrt{(E_{\tau_1} + E_{\tau_2} + \not{E}_{\mathrm{T}})^2 - (\overrightarrow{p_{\tau_1}} + \overrightarrow{p_{\tau_2}} + \overrightarrow{\not{E}_{\mathrm{T}}})^2} \qquad (13.26)$$

As can be seen from Figure 13.32, this definition successfully distinguishes between lower mass production of τ pairs and high-mass τ pairs from new massive resonant particle production.

13.8 Rejection of Events with $W^{\pm}$ and Top Quarks

W + jet(s) background becomes a dominant process because the decay of the W boson creates a very clean prompt lepton that passes all the lepton related requirements at the same rate as signal events. Moreover, jets from quarks and gluons can be identified as lepton or tau candidates. This background can be strongly suppressed by employing τ tagging criteria and event topology cuts. Tau tagging criteria is discussed in section 13.5. In this section we focus on the event topology selections used to remove W + jet(s) events.

In W + jet(s) events, unlike $X \to \tau\tau$ resonance production such as $Z' \to \tau\tau$, where the tau decay products are expected to be back-to-back in ϕ, the presence of the neutrino from the W decay and the uncorrelated jet gives rise to topologies where the jet and the lepton are not back-to-back. This topology is depicted in Figure 13.33. Therefore, one of the main discriminating variables against W + jet(s) events is the difference in ϕ between the jet and lepton directions. Figure 13.34 shows the $\cos\Delta\phi(\mu, \tau_h/jet)$ distributions for $Z \to \tau\tau$, W+jets, and a QCD muon enriched sample. Requiring e.g. $\cos\Delta\phi(\mu, \tau_h/jet) < -0.95$ is approximately 85% efficient for $Z \to \tau\tau$ and approximately 25% efficient for W+jets.

For ditau final states, the $\not\!\!E_{\rm T}$ in the event is due to the neutrinos from the tau decays and is expected to point in the direction collinear to the visible tau decay products. Furthermore, the measurement of $\not\!\!E_{\rm T}$ is completely correlated to the visible tau decay products. In W + jet(s) events, the direction and magnitude of the momentum imbalance is completely correlated to the lepton from the W boson, but uncorrelated to the jet. We require events to be consistent with this signature of a particle dacaying to two tau leptons by defining a unit vector along the bisector of visible tau decay products ($\hat{\zeta}$) and two projection variables, p_ζ and p_ζ^{vis}:

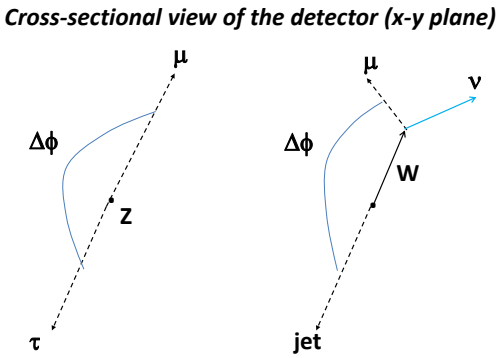

Fig. 13.33. Sketch depicting the W+jets rejection power of a $\Delta\phi$ cut.

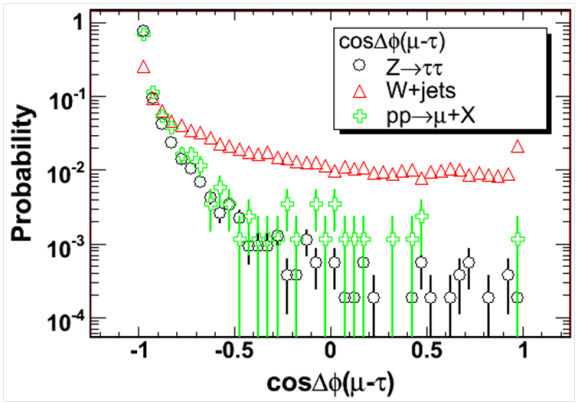

Fig. 13.34. cos$\Delta\phi$ between the muon and tau/jet directions.

$$p_\zeta^{vis} = \overrightarrow{p}_{\tau_1}^{vis}\hat{\zeta} + \overrightarrow{p}_{\tau_2}^{vis}\hat{\zeta} \tag{13.27}$$

$$p_\zeta = p_\zeta^{vis} + \overrightarrow{\not{E}_{\rm T}}\hat{\zeta} \tag{13.28}$$

The sketch in Figure 13.35 displays the definitions for p_ζ^{vis} and p_ζ. Figures 13.36 and 13.37 shows the separation between $Z' \to \tau\tau$ and W + jet(s) events in the p_ζ-p_ζ^{vis} plane. For the case of W+jets, there is no strong correlation between p_ζ^{vis} and p_ζ due to the presence of a jet that is uncorrelated to the μ and ν_μ from the W boson. However, there is a strong correlation for the case of $Z' \to \tau\tau$. To discriminate against W+jet(s) events, requirements on $\Delta\phi(\tau_1, \tau_2)$ and ζ are applied. ζ is defined as a linear combination of p_ζ and p_ζ^{vis}:

- $cos\Delta\phi(\tau_1, \tau_2) < -0.95$
- $p_\zeta - 0.875 p_\zeta^{vis} > -7$

Events containing $t\bar{t}$ contribute to the expected background in all channels containing light leptons. For $\mu\tau_h$ final states the major $t\bar{t}$ contribution comes in the form of a real light lepton from the semileptonic decay of the $W^\pm$ and a fake hadronic tau from the hadronic decay of the second $W^\pm$. These events are characterized by an isolated light lepton, passing all lepton identification and isolation requirements, accompanied by a non–isolated "hadronic tau" due to the larger multiplicity of the hadronically decaying W boson. The case where the second $W^\pm$ decays semileptonically into a tau results in an event containing isolated light leptons and taus that satisfy all identification and isolation requirements. These events are suppressed with the use of topological cuts.

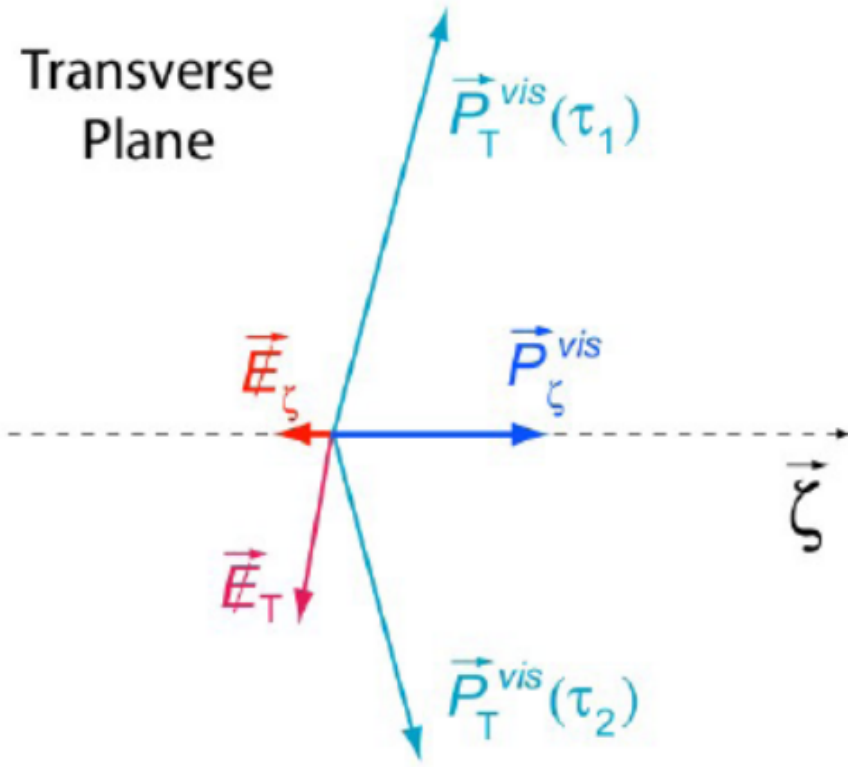

Fig. 13.35. Definitions for p_ζ and p_ζ^{vis}.

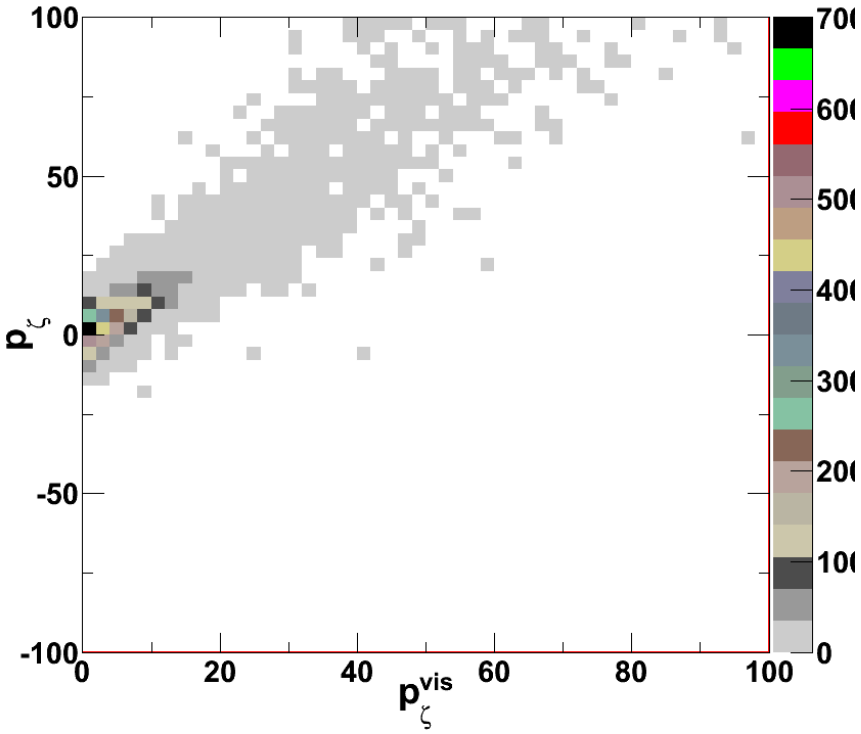

Fig. 13.36. p_ζ p_ζ^{vis} for $Z' \to \tau\tau$.

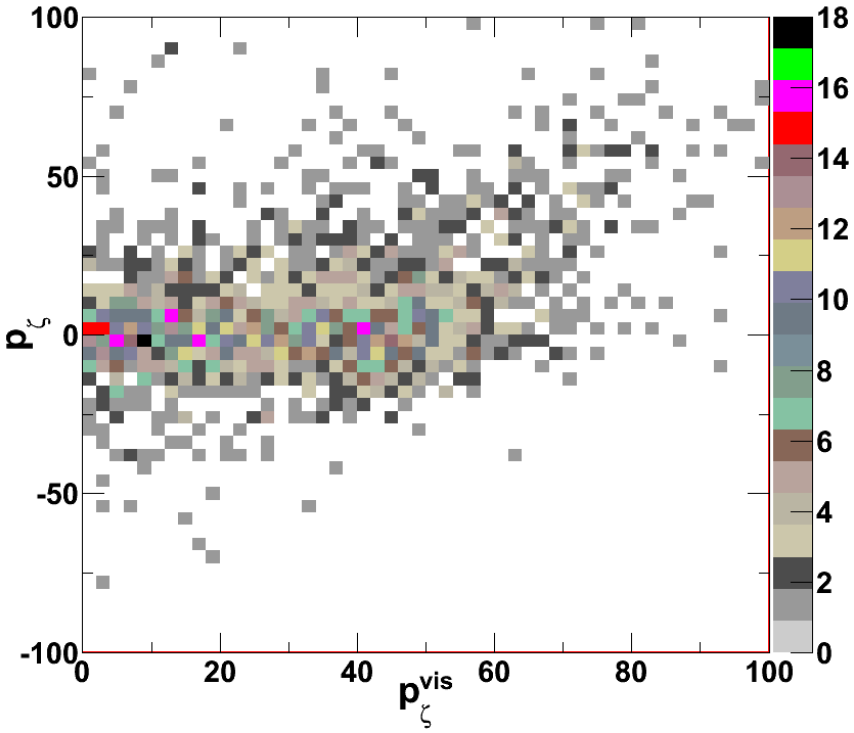

Fig. 13.37. p_ζ p_ζ^{vis} for W+Jets.

After applying lepton identification and isolation requirements, a significant background contribution from $t\bar{t}$ events remain. These events can be further suppressed with cuts that take advantage of the very different topologies between $Z' \to \tau\tau$ and $t\bar{t}$ events. The first, and most important, of these differences is the presence of b–jets in the event. Other important differences include the presence of extra jets and the angle between the $\not\!\!E_{\rm T}$ object and the highest p_T lepton in the event.

The first tool to identify and reject $t\bar{t}$ events is counting the number of jets in the event are tagged as b–jets. For our purposes, jets with $p_T > 20$ GeV/c and $|\eta| < 2.5$ are counted as b–tagged jets if the "track counting high efficiency" discriminator, described in [35], returns a value consistent with that of a b–jet. The "track counting" discriminators are very simple, yet robust discriminators that return the significance of the second (hiEff) or third (hiPurity) most significant track in the jet. In this analysis, the "medium" operating point of the "track counting high efficiency" discriminator (TCHEM) is used. The TCHEM discriminator requires a discriminant larger than 3.3 σ for a jet to be tagged as a b–jet. The mis-tag rate associated with the "medium" operating point is 1% [36]. Figure 13.38 shows the TCHE b-jet discriminator for jets in $Z' \to \tau\tau$ and various backgrounds. Jets considered for b-tagging are required to be well separated from the tagged $\mu\tau_h$ pair ($\Delta R(jet, \mu/\tau_h) > 0.5$).

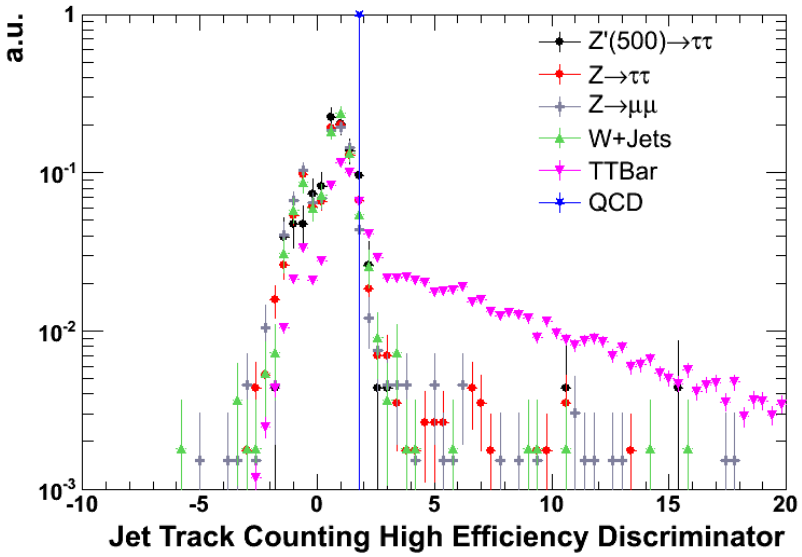

Fig. 13.38. TCHE b-jet discriminator for jets.

14. TRIGGER SELECTIONS

HLT muon candidates are reconstructed by using Level-1 muon candidates to first re-evaluate the trajectory parameters using only information from the muon chambers (Level-2). At Level-3, hits in the inner silicon tracker are combined with the hits in the muon chambers to improve the measurement of transverse momentum. It is important to note that the descoped electronics of the CSC chambers in the first muon station prevents the accurate measurement of the transverse momentum in the region $2.1 < |\eta| < 2.4$. Therefore, in order to satisfy limitations on the Level-1 trigger rate, the single muon triggers are limited to the region $|\eta| < 2.1$.

Since the trigger menus changed during the 2010 data taking period to account for increased luminosities, the highest unprescaled single muon triggers are used. Table 14.1 outlines the single muon triggers used to select events in data for various run ranges.

Table 14.1
μ triggers used for 2010 data taking.

Run Range	Single Muon Trigger
132440-147195	HLT_Mu9
147196-148058	HLT_Mu11
148059-149442	HLT_Mu15

15. EFFICIENCY MEASUREMENTS

The trigger requirement is one of the most important aspects of any analysis as it is needed to reduce the rate of collisions to a manageable rate, select events of interest, and produce little to no bias on the final analysis. Because a large part of the trigger system is based on hardware, it is often difficult to simulate trigger effects in the MC samples especially as the luminosities and thus the contributions from e.g. PU effects increase. Therefore, in cases where the trigger emulation is not included in the MC samples, the trigger efficiency must be taken from data. In cases where the trigger menu is indeed emulated in the MC samples, the trigger efficiencies extracted from data must be compared to the ones obtained from the MC samples. If efficiencies obtained from data do not agree with those extracted from MC, then proper corrections must be applied to the MC rates and efficiencies. For the case of muons, the technique used to derive the efficiencies from data is known as the *tag and probe* method.

15.1 The Tag and Probe Method

The tag and probe method makes use of the two leptons from the decays of the Z boson, $Z \to l^+l^-$, where $l = e, \mu$. The method takes advantage of two key features: (1) because two leptons are present in the event, one of the leptons can be used to significantly reduce background contamination, while the second lepton can be used to measure the efficiency of a given selection; (2) because the Z mass has been measured with good precision, the invariant mass of the two leptons can be required to be near the Z mass to obtain a clean sample of leptons. It is important to note that the method works extremely well for electrons and muons due to the very narrow mass resonance that can be attained and the low probability for jets to fake electrons and muons. However, unlike electrons and muons, a clean sample of taus cannot be attained with the tag and probe method. Therefore, in this section, focus is placed

on the use of the tag and probe method for the measurement of muon trigger and identification efficiencies. In sections to follow, the method for the extraction of tau identification efficiencies and scale factors will be discussed.

The first step in the tag and probe method is to select a sample of dimuon events. The first muon is required to pass the same muon selection criteria that is used in the analysis to ensure the measured efficiencies are not biased by factors arising from differences between selection criteria used for the tag and probe method and selection criteria used for the analysis. Furthermore, this ensures that there are no biases imposed on the second muon due to different definitions on the first muon. This muon is known as the "tag" muon. The second muon can then be used to measure the relevant efficiency. The second lepton is referred to as the "probe" muon. The number of events with at least one tag muon is n_t. In addition to the requirements on the first muon, the invariant mass of the tag and probe muons can be required to be within three standard deviations of the Z mass. At this point, even without any requirements on the probe muon, a clean sample of probe muons can be obtained. The muon trigger and identification efficiencies can be determined by counting the number of events in which the probe passed the given criteria. The number of events in which the probe muon passes the given selection is n_p. If one assumes that the probe muon is never considered as a tag muon, then the efficiency can be calculated as $\varepsilon = n_p/n_t$. However, because the probe muon can also be considered as the tag muon, the selection efficiency is instead given by

$$\varepsilon = \frac{2n_p}{n_t + n_p} \tag{15.1}$$

It is often suggested to apply a separation requirement, e.g. $\Delta R(tag, probe) > 0.5$, between the tag and probe muons to ensure that the tag and probe muons are not

double counted. However, the separation requirement can indirectly impose some bias on the quality of the muon. In this analysis, we are interested in high p_T muons ($p_T \sim 100$ GeV/c). Therefore, the efficiency can be measured as a function of p_T in order to infer the robustness of the trigger and identification requirements from the extrapolation of low p_T muons. Therefore, the selection efficiency calculation is modified to

$$\varepsilon(p_T) = \frac{2n_p(p_T)}{n_t + n_p(p_T)} \tag{15.2}$$

To test the validity of the tag and probe method, the method is first carried out using the MC simulated samples. The methodology is depicted in Figure 15.1 where the dimuon mass is displayed for tag and probe pairs passing and failing the global matching requirement between standalone muon tracks from the muon system and tracker tracks from the inner silicon tracker. Figure 15.2 shows the matching efficiency as a function of muon p_T. The tag and probe method successfully gives rise to measured efficiencies consistent with the true MC efficiencies.

15.2 Muon Efficiencies and Scale Factors

The muon trigger and identification efficiencies have been measured by another author as part of the required responsibilities of the muon object performance group (POG) within the CMS experiment [26, 27]. The MC samples do emulate the single muon triggers used in this analysis. Therefore, the measurement of the efficiencies must be used to (1) compare the efficiencies obtained from data with those obtained from MC in order to obtain a scale factor to be applied to MC rates and efficiencies,

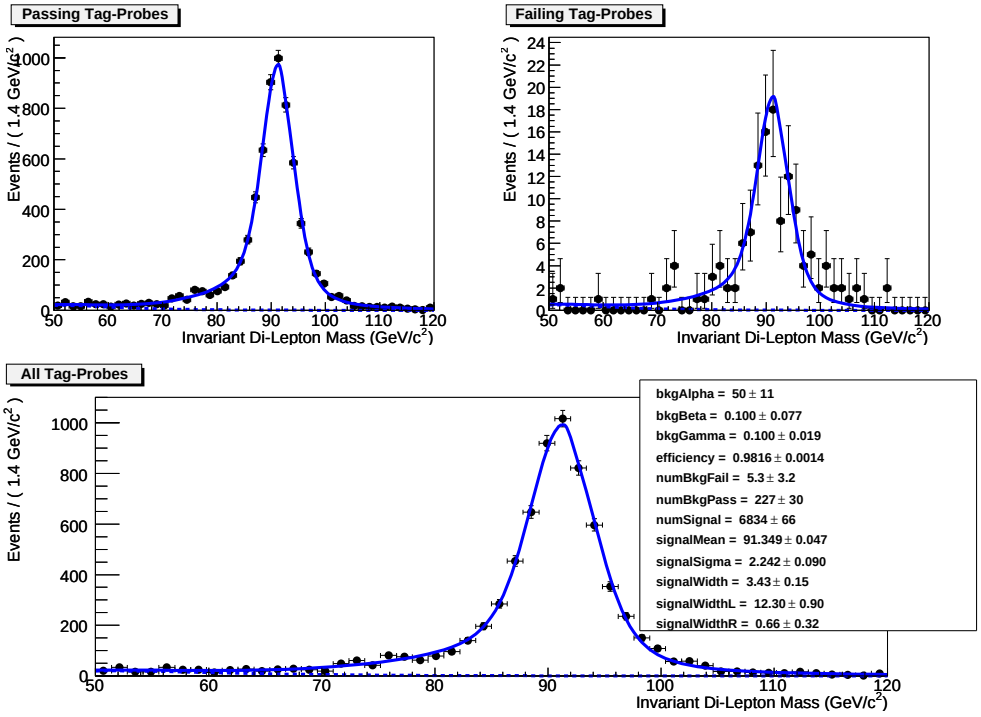

Fig. 15.1. Tag and probe dimuon mass for events that pass and fail the global muon reconstruction.

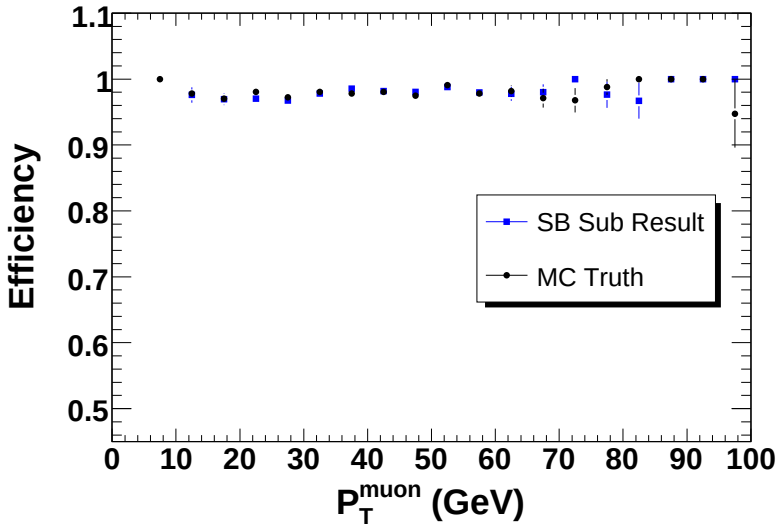

Fig. 15.2. Global muon reconstruction efficiency vs. p_T. The efficiency obtained using the tag and probe method is compared with the real MC efficiency.

and (2) assign systematic uncertainties for the imprecise knowledge of the efficiencies. The data to MC scale factor is defined as

$$f = \frac{\varepsilon_{data}}{\varepsilon_{MC}} \tag{15.3}$$

In general, it is not always true that the efficiencies and scale factors obtained by a different author using tag and probe methods can be applied to any given analysis. For example, one can make the argument that imposing the tag and probe dimuon pairs to be near the Z mass pole can unwillingly select muons with a certain quality. Furthermore, the selections used on the "tag" muon is not always the same as the selections used for a given analysis. The muon triggers and identification selections used in this analysis are the same as those used by the muon POG in the tag and probe measurements. However, to ensure that the measured scale factors can be applied in this analysis, a clean control sample of $Z \to \mu\mu$ events is obtained using the same selections used for the final analysis and rescaling the MC rates and efficiencies by the measured scale factors. If the scale factors are indeed correct and applicable for this analysis, then it is fully expected that the MC expectation and the observations from data will agree. These validations will be described in section 17.2. The overall trigger efficiencies and corresponding scale factors in the region $|\eta| < 2.1$ as measured by the CMS muon POG [26] are outlined in Table 15.1. Table 15.2 lists the muon identification efficiencies and scale factors in the region $|\eta| < 2.1$ measured by the CMS muon POG [26].

Table 15.1
μ trigger efficiencies and scale factors used for 2010 data taking (with $p_T > 20$) [26].

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$
	Efficiencies		
MC	0.9791 ± 0.0002	0.9197 ± 0.0007	0.9336 ± 0.0005
2010A	0.9523 ± 0.0079	0.7566 ± 0.0225	0.8671 ± 0.0130
2010B	0.9640 ± 0.0019	0.8841 ± 0.0051	0.9052 ± 0.0034
	Scale Factors		
2010A	0.9726 ± 0.0081	0.8227 ± 0.0244	0.9288 ± 0.0140
2010B	0.9846 ± 0.0020	0.9614 ± 0.0055	0.9696 ± 0.0037
Total	0.9837 ± 0.0019	0.9498 ± 0.0055	0.9661 ± 0.0036

Table 15.2
μ ID efficiencies used for 2010 data taking (with $p_T > 20$) [26].

	$0 < \|\eta\| < 0.8$	$0.8 < \|\eta\| < 1.2$	$1.2 < \|\eta\| < 2.1$
	Efficiencies		
MC	0.9577 ± 0.0003	0.9545 ± 0.0005	0.9622 ± 0.0004
Data	0.9530 ± 0.0047	0.9491 ± 0.0057	0.9626 ± 0.0023
	Scale Factors		
Total	0.9951 ± 0.0049	0.9944 ± 0.0060	1.0003 ± 0.0025

15.3 Tau Efficiencies and Scale Factors

Unlike muons and electrons, where a clean sample of leptons can be obtained using standard tag and probe methods, a clean sample of taus cannot be obtained using tag and probe methods. This is due to the much larger jet$\rightarrow \tau_h$ fake rate when compared to e.g. jet$\rightarrow \mu$ fake rates which is expected to be on order of 10^2 times larger. Because the tag and probe method requires the presence of two hadronically decaying taus, this implies that the background contamination will be approximately 10^2x$10^2 = 10^4$ times larger. Additionally, because taus loose energy to neutrinos, the invariant $\tau\tau$ mass does not produce a narrow peak near the Z boson mass, making it

more difficult to select a clean sample of taus without the mass window requirement.

The measurement of the tau tagging efficiencies and fake rates is a difficult task, as it requires the ability to obtain a very clean sample of real taus without the use of tag and probe methods. For a general analysis, one can obtain background free samples by using identification methods specific to that particle. However, measuring the efficiency for those identification methods requires the ability to obtain a very clean sample without using the tagging criteria to select them, thus ensuring the efficiencies and rates are not biased.

As a first step, one must choose a channel that is unbiased to the tau tagging criteria and spectra. The main difficulty in identifying hadronically decaying tau leptons is to discriminate them against jets from quarks and gluons. However, an analysis used to select $Z \to \tau\tau \to \mu\tau_h$ events may be used to obtain a clean sample of taus. The motivation behind using $Z \to \tau\tau \to \mu\tau_h$ events is that muons have the lowest jet misidentification among leptons. Therefore, the initial requirement that an event contain at least one clean muon already reduces a large amount of background contamination. Additionally, "tight" requirements can be imposed on the muon leg to drive the signal to background ratio to a point where additional bias free requirements can be used to obtain the final sample of taus.

As was described in the previous sections, the main cause for the inefficiency in isolation is due to UE/PU particles that fall into the isolation annulus and spoil the isolation requirement. The same is true for electrons and muons. This can be seen from Figures 15.3 and 15.4, which show the density of charged hadrons and photons around the τ_h/μ directions. There are some tau related effects that could cause some differences in isolation inefficiencies between muons and taus. These are related to "leakage" effects where one or more charged hadrons from the tau fall

out of the signal region and into the isolation annulus. However, in this analysis, τ_h's are required have only one charged hadron in the signal region. Therefore, the analysis is not sensitive to "leakage" effects for three pronged taus. There is also a leakage effect due to $\pi^{0\prime}s$ that decay to photons and then subsequently convert within the tracker material and cause the electron-positron pairs to be swept by the magnetic field into the isolation annulus. However, the Ecal signal cone used in this analysis is specifically designed to recover the efficiency loss due to photon conversions. Therefore, it is safe to say that for the tau selections used in this analysis, the inefficiencies in isolation for taus and muons are expected to be similar. Therefore, Drell-Yan $pp \rightarrow \gamma^*/Z \rightarrow \mu\mu$ events using standard tag and probe methods can be used to measure the isolation efficiencies for τ_h's. Since isolation efficiencies can be measured elsewhere, the isolation requirement on the τ_h leg can be used to reduce the background contamination to a level where the other tau identification efficiencies can be measured.

As was the case for muons, measurement of the τ_h identification efficiencies and scale factors is the responsibility of the tau performance object group (POG). The tau identification scale factor as measured by the CMS Tau POG [37] is

$$f_{\tau_h} = 1.01 \pm 0.26 \tag{15.4}$$

Because obtaining a clean sample of taus using the method presented in this section requires extremely tight requirements in order to reduce the background contamination to a negligible level without the use of tau tagging criteria, statistics are limited with only 36.15 pb^{-1} of data. Therefore, scale factors cannot be determined as a function of p_T or η. However, following the same methodology used for muons, the use of this scale factor can be validated by obtaining a clean control sample of

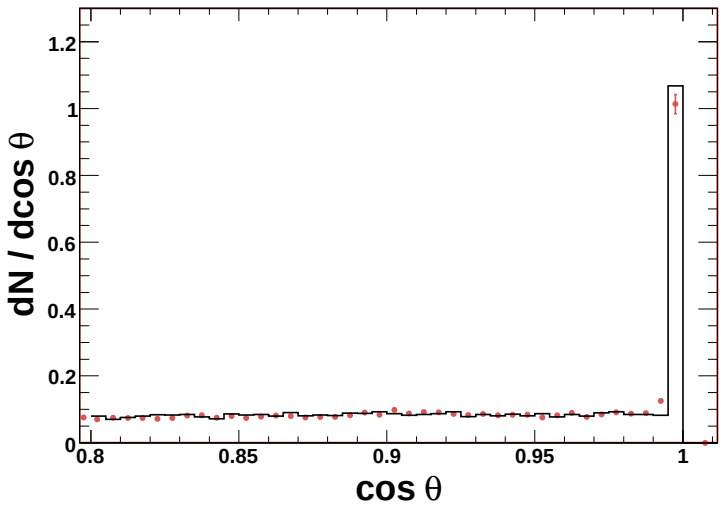

Fig. 15.3. Density of PF charged hadrons with respect to the τ_h/μ directions for muons from $Z \to \mu\mu$ (solid black line) and τ_h's from $Z \to \tau\tau$ (peach dots).

$Z \to \tau\tau$ events obtained using the same selections used for the final $X \to \tau\tau$ analysis and rescaling the MC rates and efficiencies by the measured scale factor. If the scale factors are indeed correct and applicable for this analysis, then it is fully expected that the MC expectation and the observations from data will agree. These validations will be described in section 17.5.

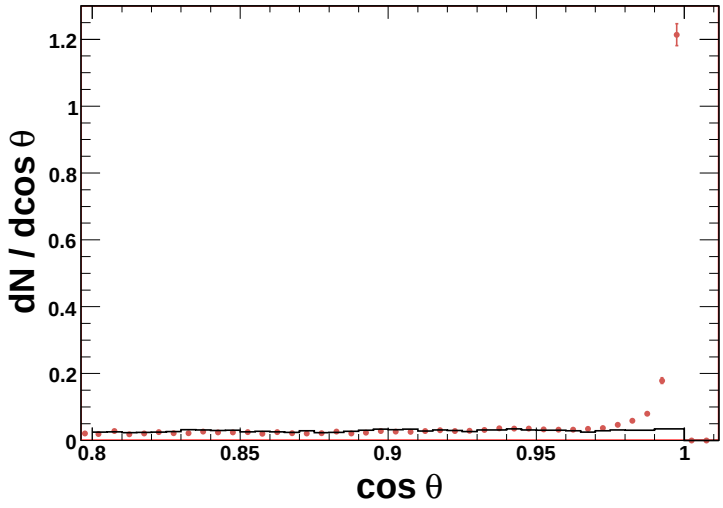

Fig. 15.4. Density of PF photons with respect to the τ_h/μ directions for muons from $Z \to \mu\mu$ (solid black line) and τ_h's from $Z \to \tau\tau$ (peach dots).

16. ANALYSIS SELECTIONS AND EFFICIENCIES

The motivation for analyzing events where one tau lepton decays to a muon, while the other decays to hadrons is the same for all $\tau\tau$ related analyses. Because muons have the lowest jet misidentification among leptons, the mere requirement of a muon removes a substantial amount of background processes. Once this requirement is made, the main source of background for many related searches is due to Drell-Yan processes giving rise to tau leptons. Figure 4.1 shows the Feynman diagram for Drell-Yan $pp \to \gamma^*/Z \to \tau\tau$ production. Because we seek particles with masses much larger than that of the Z boson, this source of background can be easily discriminated against by looking at larger reconstructed $\tau\tau$ mass regions. This process, however, can serve as a control sample to validate the τ_h identification criteria and ensure the robustness of the analysis, especially for high p_T taus. Other main sources of background include (1) QCD events where b jets produce muons associated to jets ($b \to \nu\mu c$), (2) W + n jet events where the W boson decays to a muon and a jet is misidentified as a hadronically decaying tau, and (3) $t\bar{t}$ events where two leptons can come from the prompt decay of W bosons or one mis-identified tau from a jet. The requirements used to select $\mu\tau_h$ pairs is factorized in to four categories: *acceptance*, μ *identification*, τ *identification*, and *topological* cuts. Acceptance criteria is completely driven by the limitations of the CMS detector and the need to maximize analysis sensitivity while also minimizing systematic effects. For example, as discussed in section 14, limitations on the Level-1 trigger rate force us to select μ's in the region $|\eta| < 2.1$. Additionally, track isolation can be difficult to understand in the region $|\eta| > 2.1$ where the isolation regions extend beyond the edge of the silicon tracker. Similarly, in order to minimize systematic effects, the p_T thresholds on the ditau candidates are chosen such that they fall on the plateu of the trigger turn-on curves. Although it is possible to achieve slightly better sensitivity by increasing the thresholds, the selections are also driven by the need to obtain a $Z \to \tau\tau$ control sample

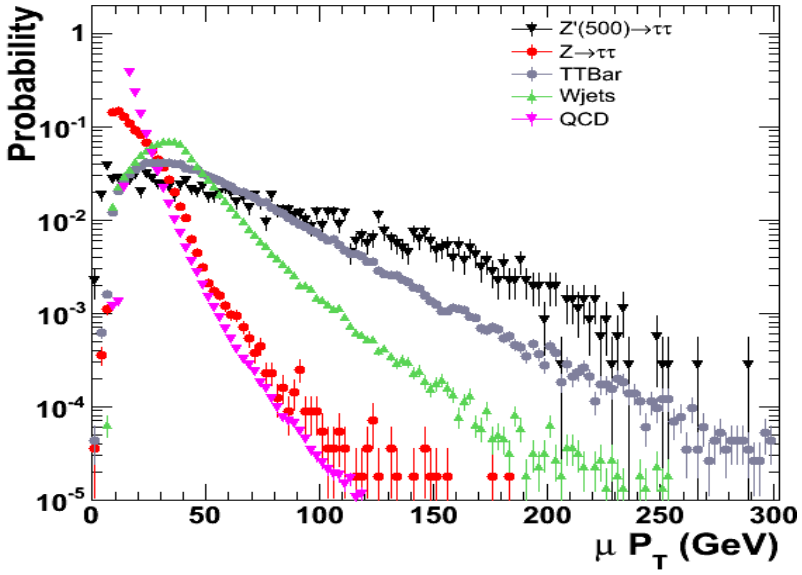

Fig. 16.1. μ p_T distributions for signal and various backgrounds.

with minimal modifications to the final selection criteria. Figures 16.1 and 16.2 show the p_T distributions for signal and background processes relevant to this analysis. As discussed in section 13.2, the muon identification criteria is designed mostly to discriminate against cosmic muons, punch-through pions, and muons associated to jets from decays in flight. Tau identification is described in section 13.5 and is designed to discriminate against hadronic jets produced from the fragmentation of quarks and/or gluons. Finally, topological cuts are mainly used to minimize the remaining W+jet(s) and $t\bar{t}$ contributions that remain after muon and tau identification criteria has been imposed (see section 13.8). The entire signal selection criteria are summarized below.

Acceptance Selection:

- ≥ 1 Global μ with $|\eta| < 2.1, p_T > 20$ GeV/c
- ≥ 1 PF τ with $|\eta| < 2.1, p_T > 20$ GeV/c, and leading track with $p_T > 5$ GeV/c

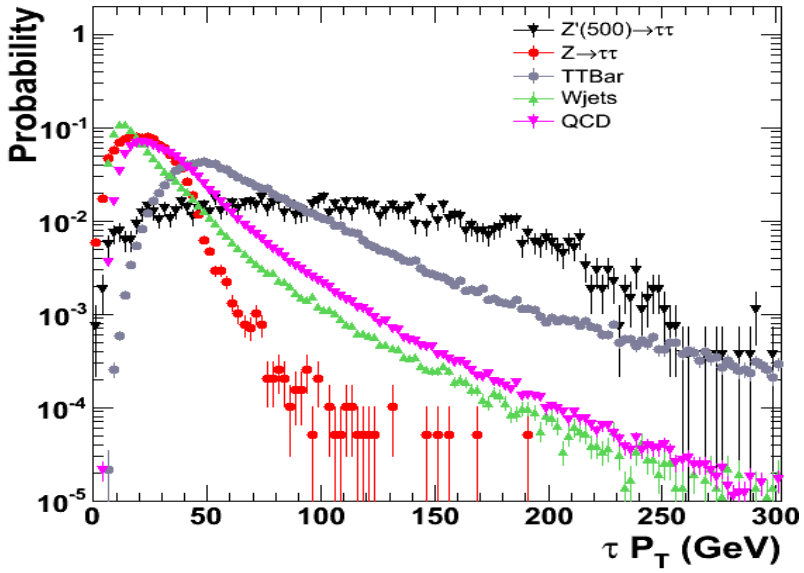

Fig. 16.2. τ_h/jet p_T distributions for signal and various backgrounds.

- $\Delta R(\mu, \tau_h) > 0.7$

μ Identification:

- To remove cosmic muons, the μ impact parameter must satisfy: $|d_0| < 0.2$ cm
- To remove punch-through pions, the π veto must be satisfied: $0.8*C+b*S > 1$
- Track Isolation: $\sum p_T^{trk} < 1.0$ GeV/c ($p_T^{trk} > 0.7$ GeV/c, $\Delta R_{veto} = 0.01$, $\Delta R_{iso} = 0.4$)
- Ecal Isolation : $\sum E_T^{ecal} < 1.0$ GeV/c ($E_T^{Ecal\ RecHit} > 0.3$ GeV/c, $\Delta R_{veto} = 0.01$, $\Delta R_{iso} = 0.4$)

τ_h Identification:

- Muon veto: $0.8 * C + 1.2 * S < 1$
- Exactly 1 signal charged hadron ($p_T^{chad} > 1$ GeV, $\Delta R_{sig} = 5/E_T$)
- Track Isolation : $\sum p_T^{chad} < 1.0$ GeV ($p_T^{chad} > 1$ GeV/c, $\Delta R_{iso} = 0.5$, $\Delta R_{Track\ Signal} = 5/E_T$)
- Ecal Isolation : $\sum E_T^{\gamma' s} < 1.0$ GeV ($E_T^{\gamma' s} > 1.0$ GeV, $R_\eta = 0.07$, $R_\phi = 0.15$)

Topological requirements:

- $\cos \Delta\phi(\mu, \tau_h) <$ -0.95. The τ_h jet direction is calculated using the sum of the four-momenta of signal cone constituents:

$$\overrightarrow{p}_{\tau_h} = \sum_i \overrightarrow{p}^{\,i}_{signal\ constituents} \tag{16.1}$$

- $Q(\mu) \times Q(\tau_h) < 0$ (τ charge is defined as the charge of the leading track)
- $\not\!\!E_{\rm T} > 30$ GeV

- $P_\zeta - 0.875 \times P_\zeta^{vis} > -7$
- 0 jets tagged as b-jets

Table 16.1 shows the relative $\mu\tau_h$ selection efficiencies for various Z' masses, while Table 16.2 shows similar efficiencies for the relevant SM backgrounds. The relative efficiencies are defined with respect to the previous selections as follows:

$$\varepsilon_{rel} = \frac{\text{Number of Events Passing Cut } i+1}{\text{Number of Events Passing Cut } i} \tag{16.2}$$

The overall signal selection efficiency

$$\varepsilon_{signal} = \frac{\text{Number of Events Passing All Selections}}{\text{Number of Events Before Any Selections}} \tag{16.3}$$

is summarized in Table 16.3. Therefore, for a luminosity of 36.15 pb^{-1}, a Z' mass of 350 GeV/c^2 is expected to produce $N = \sigma \cdot L \cdot \varepsilon \cdot BR(\tau\tau \to \mu\tau_h) = 8.06 \cdot 36.15 \cdot 0.1154 \cdot 0.225 \sim 7.5$ events. Similarly, for a luminosity of 36.15 pb^{-1}, $t\bar{t}$ is expected to contribute $N = \sigma \cdot L \cdot \varepsilon \sim 150 \cdot 36.15 \cdot 0.0001 \sim 0.5$ events. The uncertainties on the efficiencies listed in Table 16.3 are entirely determined by the statistics of the MC simulated samples.

Table 16.1
Relative Cut Efficiencies for Various Z' Masses.

Cut/Selection	Z'(350)$\rightarrow \tau\tau$	Z'(400)$\rightarrow \tau\tau$	Z'(500)$\rightarrow \tau\tau$	Z'(600)$\rightarrow \tau\tau$	Z'(700)$\rightarrow \tau\tau$
μ $p_T > 20$	74.85 ± 0.71	78.06 ± 0.70	81.76 ± 0.63	84.48 ± 0.58	86.19 ± 0.55
μ $\lvert\eta\rvert < 2.1$	93.46 ± 0.47	93.81 ± 0.46	94.78 ± 0.40	95.37 ± 0.37	96.66 ± 0.31
τ $p_T > 20$	94.38 ± 0.45	95.29 ± 0.42	96.53 ± 0.34	97.67 ± 0.27	97.93 ± 0.25
τ $\lvert\eta\rvert < 2.1$	84.96 ± 0.72	86.30 ± 0.69	87.24 ± 0.63	88.49 ± 0.58	90.04 ± 0.53
τ Seed $p_T > 5$	92.18 ± 0.59	93.10 ± 0.55	92.77 ± 0.52	92.50 ± 0.51	93.31 ± 0.47
μ $\lvert d_0\rvert < 0.2$	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00	100.0 ± 0.00
μ π veto	98.91 ± 0.24	99.14 ± 0.21	98.35 ± 0.27	98.88 ± 0.21	99.14 ± 0.18
μ ECAL Iso. $\sum p_T < 1$	89.48 ± 0.70	89.83 ± 0.68	88.53 ± 0.67	88.28 ± 0.65	87.89 ± 0.63
μ Track Iso. $\sum p_T < 1$	88.83 ± 0.76	87.43 ± 0.79	87.55 ± 0.74	88.28 ± 0.69	87.73 ± 0.68
τ μ veto	98.95 ± 0.26	99.03 ± 0.25	98.63 ± 0.28	98.13 ± 0.31	97.74 ± 0.33
τ 1 prong	72.65 ± 1.15	74.04 ± 1.12	74.64 ± 1.05	75.58 ± 0.99	76.24 ± 0.95
τ ECAL Iso. $\sum p_T < 1$	77.56 ± 1.26	77.12 ± 1.25	75.33 ± 1.20	76.43 ± 1.12	74.18 ± 1.12
τ Track Iso. $\sum p_T < 1$	85.71 ± 1.20	85.11 ± 1.21	85.17 ± 1.14	85.64 ± 1.06	84.90 ± 1.07
$\cos\Delta\phi(\mu,\tau) < -0.95$	88.15 ± 1.20	90.58 ± 1.07	91.17 ± 0.99	93.27 ± 0.82	94.67 ± 0.73
$Q(\mu) * Q(\tau jet) < 0$	97.97 ± 0.56	98.81 ± 0.42	97.21 ± 0.60	98.17 ± 0.45	96.69 ± 0.60
$\not\!\!E_{\mathrm{T}} > 30$	62.84 ± 1.93	69.32 ± 1.79	75.03 ± 1.60	78.76 ± 1.40	82.51 ± 1.28
$P_\zeta - 0.875 P_\zeta^{vis} > -7$	88.32 ± 1.62	91.32 ± 1.31	92.18 ± 1.15	91.56 ± 1.07	92.80 ± 0.96
0 b-tagged jets	100.0 ± 0.00	99.76 ± 0.24	100.0 ± 0.00	99.51 ± 0.28	99.25 ± 0.33

Table 16.2
Relative Cut Efficiencies for SM Backgrounds.

Cut/Selection	Z$\rightarrow \tau\tau$	Z$\rightarrow \mu\mu$	W+Jets	QCD	$t\bar{t}$
μ $p_T > 20$	39.57 ± 0.14	73.30 ± 0.04	80.49 ± 0.05	11.67 ± 0.02	63.07 ± 0.08
μ $\lvert\eta\rvert < 2.1$	91.17 ± 0.13	95.66 ± 0.02	89.33 ± 0.04	94.00 ± 0.04	96.91 ± 0.03
τ $p_T > 20$	64.50 ± 0.22	86.27 ± 0.04	49.73 ± 0.07	73.99 ± 0.07	98.91 ± 0.02
τ $\lvert\eta\rvert < 2.1$	90.60 ± 0.17	85.81 ± 0.04	89.89 ± 0.06	84.62 ± 0.07	98.78 ± 0.02
τ Seed $p_T > 5$	90.83 ± 0.18	97.06 ± 0.02	88.62 ± 0.06	85.82 ± 0.07	96.07 ± 0.04
μ $\lvert d_0\rvert < 0.2$	99.98 ± 0.01	100.00 ± 0.00	99.98 ± 0.01	99.86 ± 0.01	99.91 ± 0.01
μ π veto	95.69 ± 0.13	99.80 ± 0.01	97.96 ± 0.03	88.71 ± 0.07	93.95 ± 0.05
μ ECAL Iso. $\sum p_T < 1$	94.22 ± 0.15	98.11 ± 0.02	91.50 ± 0.06	8.09 ± 0.07	62.51 ± 0.10
μ Track Iso. $\sum p_T < 1$	89.82 ± 0.20	95.95 ± 0.03	88.08 ± 0.07	10.64 ± 0.26	81.60 ± 0.11
τ μ veto	94.21 ± 0.17	15.00 ± 0.05	99.22 ± 0.02	98.22 ± 0.34	97.32 ± 0.05
τ 1 prong	55.23 ± 0.36	21.75 ± 0.15	17.12 ± 0.09	15.53 ± 0.94	36.46 ± 0.15
τ ECAL Iso. $\sum p_T < 1$	71.29 ± 0.45	47.03 ± 0.38	29.33 ± 0.27	27.71 ± 2.95	38.11 ± 0.25
τ Track Iso. $\sum p_T < 1$	87.06 ± 0.39	58.31 ± 0.55	26.93 ± 0.48	36.51 ± 6.07	59.80 ± 0.40
$\cos\Delta\phi(\mu,\tau) < -0.95$	69.51 ± 0.58	60.77 ± 0.71	21.29 ± 0.85	56.52 ± 10.34	15.55 ± 0.39
$Q(\mu) * Q(\tau jet) < 0$	99.45 ± 0.11	92.99 ± 0.48	86.10 ± 1.65	76.92 ± 11.69	97.16 ± 0.45
$\not\!\!E_{\mathrm{T}} > 30$	3.97 ± 0.29	0.6785 ± 0.16	28.19 ± 2.21	20.00 ± 12.65	83.60 ± 1.01
$P_\zeta - 0.875 P_\zeta^{vis} > -7$	89.14 ± 2.35	55.56 ± 11.71	48.72 ± 4.62	50.00 ± 35.36	54.03 ± 1.49
0 b-tagged jets	98.72 ± 0.90	100.00 ± 0.00	100.00 ± 0.00	100.00 ± 0.00	25.87 ± 1.78

Table 16.3
Signal Acceptance for Various Masses.

Z' Mass (GeV/c^2)	Acceptance
350	0.0813 ± 0.0041
400	0.1025 ± 0.0047
500	0.1154 ± 0.0050
600	0.1418 ± 0.0054
700	0.1595 ± 0.0058
800	0.1673 ± 0.0059
900	0.1690 ± 0.0060
1000	0.1785 ± 0.0067

17. BACKGROUND ESTIMATIONS AND VALIDATIONS

For tau related analyses, it is generally not expected that predictions based on the MC simulated samples will agree with the observations from real *pp* collisions. Backgrounds relevant to any tau related analysis are dominated by the probabilities for jets from the fragmentation of quarks and/or gluons to produce signatures that can look like those of hadronically decaying taus. The fragmentation of quarks and gluons is a difficult process to model in simulated samples due to our inability to probe quarks and gluons as free particles. Furthermore, additional effects such as the modeling of pile-up and material budget are all important contributors to tau identification and jet$\rightarrow \tau_h$ fake rates. Since the MC simulated samples are usually not reliable, this analysis makes use of data-driven methods to extract the background contributions in the signal region. This is accomplished by creating background enriched control samples obtained by modifying the final selection criteria. The modifications imposed to obtain background enriched regions must not bias the shapes and the ability to measure selection efficiencies. Once a background enhanced region is created, the selection efficiencies are measured in those regions and used to extrapolate to the region where a new high mass resonance is expected.

17.1 QCD Background Estimation from Data

To obtain a statistically significant sample of events, the selection criteria outlined in section 16 can be relaxed to increase the contribution from each SM background:

- "Loose" μ track isolation : $0 <$ μ Track Isolation < 15
- "Loose" τ track isolation : $0 <$ τ_h Track Isolation < 15
- Remove the requirement on $Q(\mu) * Q(\tau_{seed})$
- Remove the requirement on $\not\!\!E_{\mathrm{T}}$

- Remove the requirement on ζ

Figure 17.1 shows the μ track isolation distribution for events passing the criteria defined above. Figure 17.2 will be discussed in the sections to follow. One can see that, in addition to the above criteria, a pure sample of QCD events can be obtained by applying an "anti-isolation" cut on the μ leg ($4 <$ μ Track Isolation < 15). This defines the first control sample of QCD events, QCD Region 1:

- Begin with the selection criteria outlined in section 16
- Modify μ Isolation: $4 <$ μ Track Isolation < 15
- Modify τ Isolation: $0 <$ τ Track Isolation < 15
- Remove the requirement on $Q(\mu) * Q(\tau_{seed})$
- Remove the requirement on $\not\!\!E_{\mathrm{T}}$
- Remove the requirement on ζ

Table 17.1 shows the number of observed events in data as well as the number of expected events in MC, while Figures 17.3 – 17.6 show the distributions obtained from QCD Region 1. The uncertainties in the expected rates are determined by the available statistics of the simulated samples. The purity of QCD events in this region, QCD Region 1, is $> 99\%$. This control sample, QCD Region 1, can be used to measure efficiencies and extract the number of QCD events in the signal region. It is important to note that agreement between the MC expectations and shapes are not required. In fact, disagreement is usually expected, especially for backgrounds such as QCD, due to the usual inability for the MC simulated samples to properly describe the data in cases where the probability for a jet to fake a tau due to the hadronization of quarks and/or gluons must be modeled correctly. However, Figures 17.3 – 17.6 do indeed show agreement, even for $\not\!\!E_{\mathrm{T}}$ related variables where several effects such as pile-up must be well modeled. The fact that $\not\!\!E_{\mathrm{T}}$ related variables do show agreement

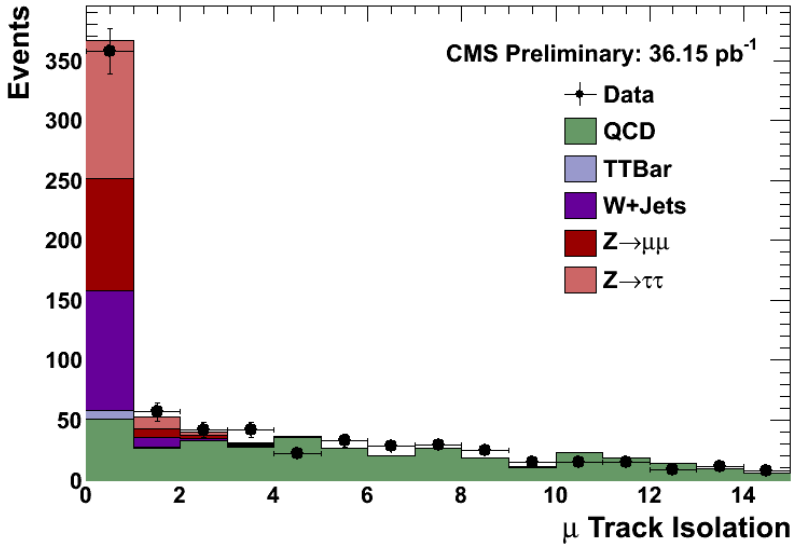

Fig. 17.1. μ track isolation distribution for events with μ track isolation < 15, τ track isolation < 15, No $\not\!\!E_{\mathrm{T}}$, No ζ, and No OSLS requirements.

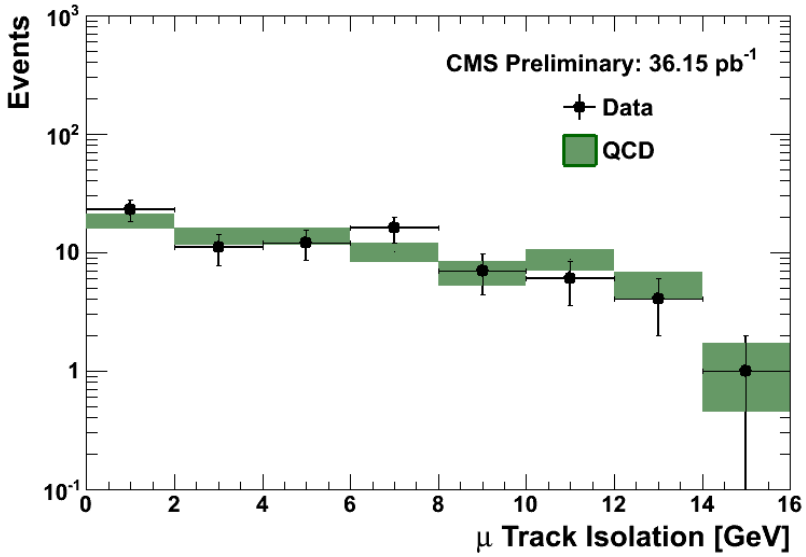

Fig. 17.2. μ track isolation distribution for events in QCD Region 2.

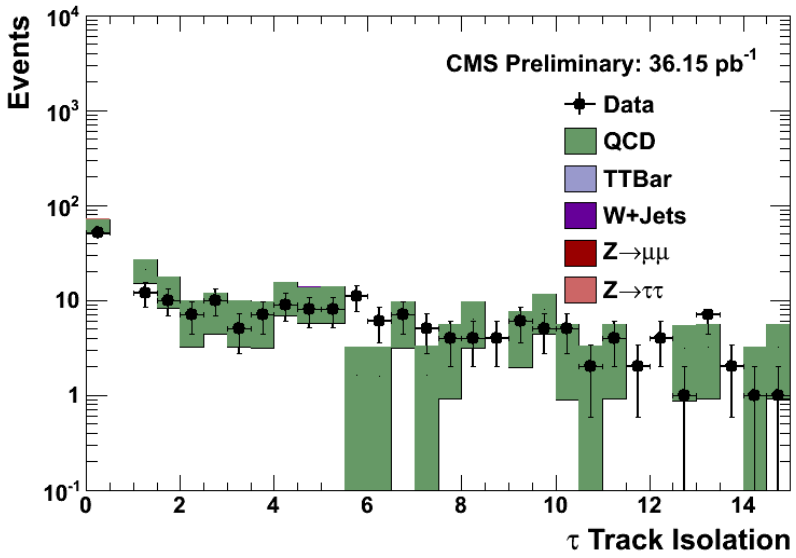

Fig. 17.3. τ_h track isolation distribution from events in QCD Region 1.

in this region gives extreme confidence that the corrections for pile-up have been properly modeled.

Table 17.1
Events in data and MC from QCD Region 1.

Sample	Events
Data	208
QCD	194.3 ± 17.52
$W + Jets$	0.98 ± 0.26
$t\bar{t}$	0.12 ± 0.02
$Z \to \tau\tau$	1.10 ± 0.18
$Z \to \mu\mu$	0.68 ± 0.13

The method of using the background enhanced region to measure selection efficiencies to be used to extrapolate to the signal region will only work if the requirement of a non-isolated μ does not bias the ability to measure efficiencies and extract correct

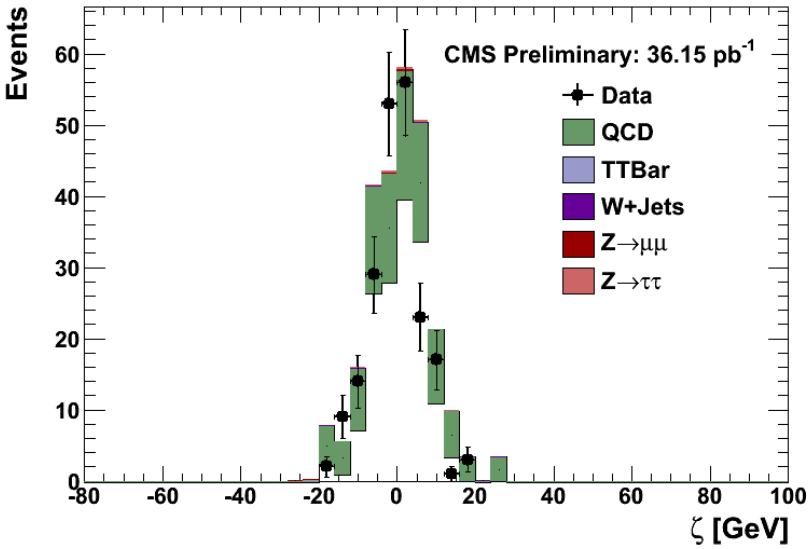

Fig. 17.4. ζ distribution from events in QCD Region 1.

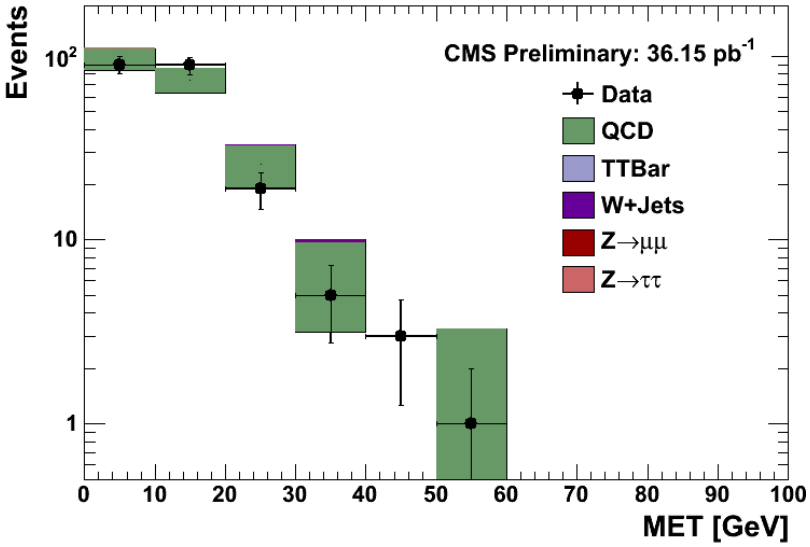

Fig. 17.5. $\not\!\!E_{\mathrm{T}}$ distribution from events in QCD Region 1.

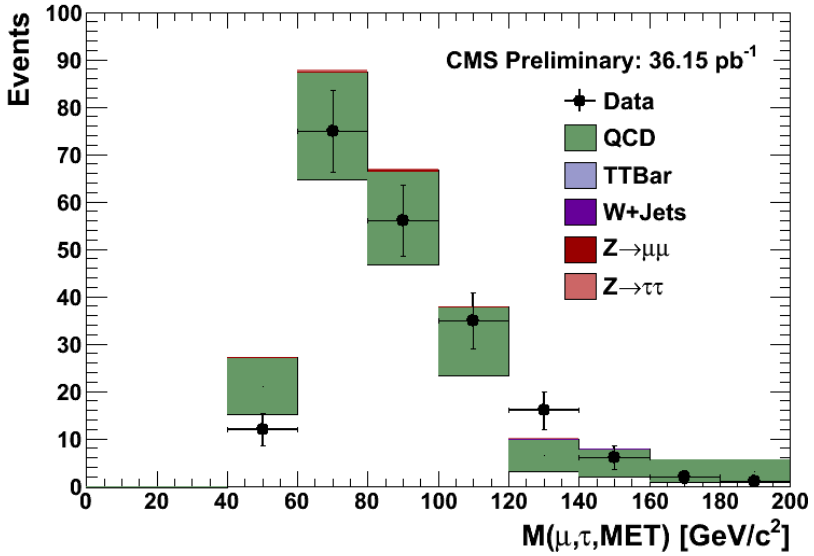

Fig. 17.6. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ distribution from events in QCD Region 1.

mass shapes. To check this is a safe requirement, the study was carried out using the MC simulated samples to determine whether μ isolation was strongly correlated to τ_h isolation and $\not\!\!E_{\mathrm{T}}$ related quantities. Figures 17.7 and 17.8 show the distributions for τ_h track isolation and ζ with "loose" ($4 <$ μ Track Isolation < 15) and "tight" (μ Track Isolation < 4) isolation requirements on the μ candidate. One can see that requiring a non-isolated μ does not bias the shapes and efficiencies. Therefore, the control sample from QCD Region 1 can be used to measure ε^{τ} Track Iso, $\varepsilon^{\zeta + \not\!\!E_{\mathrm{T}}}$, and $\varepsilon^{Q(\mu)*Q(\tau_{seed})}$:

$$\varepsilon^{\tau} \text{ Track Iso} = \frac{\text{Number of Events Passing } \tau_h \text{ Track Isolation}}{\text{Number of Events in QCD Region 1}} \tag{17.1}$$

$$\varepsilon^{\zeta + \not\!\!E_{\mathrm{T}}} = \frac{\text{Number of Events Passing } \zeta \text{ and } \not\!\!E_{\mathrm{T}}}{\text{Number of Events in QCD Region 1}} \tag{17.2}$$

$$\varepsilon^{Q(\mu)*Q(\tau_{seed})} = \frac{\text{Number of Events Passing } Q(\mu) * Q(\tau_{seed})}{\text{Number of Events in QCD Region 1}} \tag{17.3}$$

To complete the data-driven QCD background estimation, a second QCD control sample, QCD Region 2, is required that will allow for the measurement the μ track isolation efficiency:

- Begin with the selection criteria outlined in section 16
- Modify μ Isolation: $0 <$ μ Track Isolation < 15
- Modify τ Isolation: $1 <$ τ Track Isolation < 15
- Modify the requirement on $Q(\mu) * Q(\tau_{seed})$: $Q(\mu) * Q(\tau_{seed}) > 0$

- Remove the requirement on $\not\!\!E_{\mathrm{T}}$
- Remove the requirement on ζ
- Apply a requirement on the μ-$\not\!\!E_{\mathrm{T}}$ transverse mass: $M_T(\mu, \not\!\!E_{\mathrm{T}}) < 10$

Although the requirement on the μ-$\not\!\!E_{\mathrm{T}}$ transverse mass, $M_T(\mu, \not\!\!E_{\mathrm{T}}) < 10$, is not used in the final selection, it is used to select events in QCD Region 2 and ensure the contamination from W+jets is negligible. Table 17.2 shows the number of events in data and MC for QCD Region 2, while Table 17.3 shows the measurement of the efficiencies from data and the corresponding efficiencies for MC (for comparison only). Figure 17.2 shows the μ isolation distribution for events in QCD Region 2. As was the case for QCD Region 1, this region has a purity $> 99\%$ and shows very good agreement between MC and data. Having measured the relevant efficiencies, the number of QCD events in the signal region can be extracted as follows:

$$N_{QCD}^{signal} = \underbrace{N_{QCD}^{pure} \varepsilon^{\tau TrkIso} \varepsilon^{MET,\ \zeta} \varepsilon^{SSOS}}_{QCD\ region\ 1} \underbrace{\frac{N_{\mu\ \text{Track Iso}}^{[0,1]}}{N_{\mu\ \text{Track Iso}}^{[4,15]}}}_{QCD\ region\ 2} \tag{17.4}$$

Table 17.3 shows the expected number of QCD events in the $Z' \to \tau\tau$ signal region extracted in a data-driven way using the above methodology. The expected QCD contribution is 0.205 ± 0.101, validating the use of the $\mu\tau_h$ final state and one prong taus to minimize the contribution from a background that is generally the largest source of contamination for tau related analyses.

Table 17.2
Events in data and MC from QCD Region 2.

Sample	Events
Data	79
QCD	77.95 ± 12.49

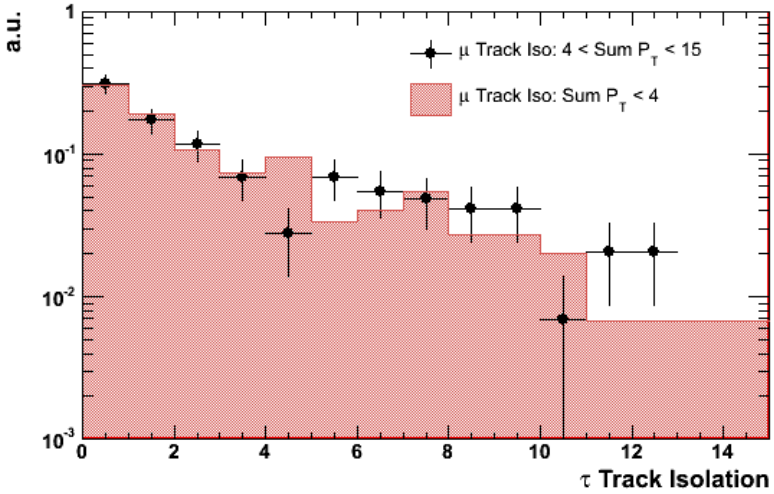

Fig. 17.7. τ_h track isolation distributions with "loose" and "tight" μ isolation requirements.

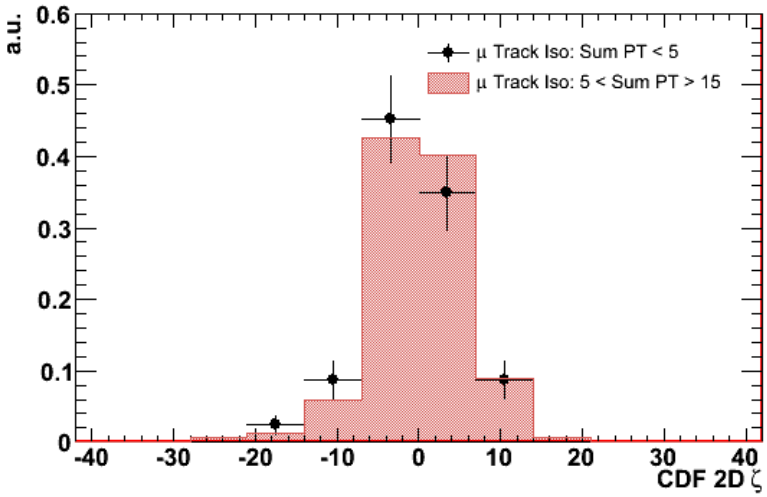

Fig. 17.8. ζ distributions with "loose" and "tight" μ isolation requirements.

Table 17.3
QCD Extraction Efficiencies.

Cut	Data	MC
ε^{τ} Track Iso	0.245 ± 0.030	0.267 ± 0.041
$\varepsilon^{\zeta} + \not{E}_{\mathrm{T}}$	0.029 ± 0.012	0.031 ± 0.016
$\varepsilon^{Q(\mu)*Q(\tau_{seed})<0}$	0.457 ± 0.035	0.527 ± 0.045
ε^{μ} Track Iso	0.177 ± 0.043	0.153 ± 0.025
$N_{\mu}^{[0,1]}$ Track Iso$/N_{\mu}^{[4,15]}$ Track Iso	0.304 ± 0.068	0.260 ± 0.040
Expected Number of Events	0.205 ± 0.101	1.579 ± 1.579

17.2 $Z \to \mu\mu$ Background Estimation from Data

Although it is believed that the signal contamination from $Z \to \mu\mu$ is negligible due to the low probability for a muon to fake a hadronically decaying tau (10^{-4} –10^{-3}), the fact that muons have the lowest jet misidentification among leptons can be used to select a clean sample of $Z \to \mu\mu$. The pure sample of $Z \to \mu\mu$ can be used to (1) measure selection efficiencies and estimate the signal contribution, and (2) validate the muon identification criteria. Using the final analysis selections as a starting point, a clean sample of $Z \to \mu\mu$ can be obtained by noting that the main discriminator against $Z \to \mu\mu$ events is $\not{E}_{\mathrm{T}}$. Measurement of non-zero $\not{E}_{\mathrm{T}}$ values are mostly driven by the mis-measurement of the μ momentum. Furthermore, because the probability for a charged pion to fake a μ is of the order of 10^{-3}, dimuon events are highly suppressed by the requirement of the μ veto on the τ_h leg. Therefore, to obtain a pure sample of $Z \to \mu\mu$ events, the final selection criteria described in section 16 is modified as follows:

- Remove the requirement on $\not{E}_{\mathrm{T}}$
- Replace the μ veto cut on the tau leg with an anti-muon veto requirement: $0.8 * C + 1.2 * S > 1$ (select PF τ's that are "muon-like")

Figure 17.9 shows the μ-τ-$\not\!\!E_{\mathrm{T}}$ invariant mass in the $Z \to \mu\mu$ control region defined above, while Table 17.4 shows the number of observed events in data and expected events in MC.

Table 17.4
Events in the $Z \to \mu\mu$ control region for data and MC.

Sample	Events
Data	5898
QCD	0.00 ± 1.57
$W + Jets$	0.28 ± 0.14
$t\bar{t}$	0.50 ± 0.05
$Z \to \tau\tau$	16.8 ± 0.70
$Z \to \mu\mu$	5765 ± 11.0

The number of $Z \to \mu\mu$ events in the signal region can be estimated as follows:

$$N_{Z\to\mu\mu}^{signal} = N_{Z\to\mu\mu}^{pure} \varepsilon_{\not\!\!E_{\mathrm{T}}} \varepsilon_{\mu \text{ veto}} \tag{17.5}$$

where $\varepsilon_{\not\!\!E_{\mathrm{T}}}$ is the $Z \to \mu\mu$ efficiency for the requirement $\not\!\!E_{\mathrm{T}} > 30$ and $\varepsilon_{\mu \text{ veto}}$ is the probability for a μ to fake a pion as measured in [26, 27] using the tag and probe method described in section 15. Note that since $\varepsilon_{\not\!\!E_{\mathrm{T}}}$ and $\varepsilon_{\mu \text{ veto}}$ are uncorrelated, they can be measured independently in the $Z \to \mu\mu$ control region. The ability to measure selection efficiencies independently from other selections allows for a more precise estimation of the background contributions in the signal region. Table 17.5 shows the measurement of the efficiencies from data, and the expected contribution in the $Z' \to \tau\tau$ signal region using the efficiencies listed in Table 17.5 and eq. 17.5. As expected, because the $\mu \to \tau_h$ fake rate is small (10^{-4} –10^{-3}), this background is a small contribution to the signal region (expected 0.225 ± 0.054). All trigger and efficiency scale factors described in section 15 have been applied in the numbers

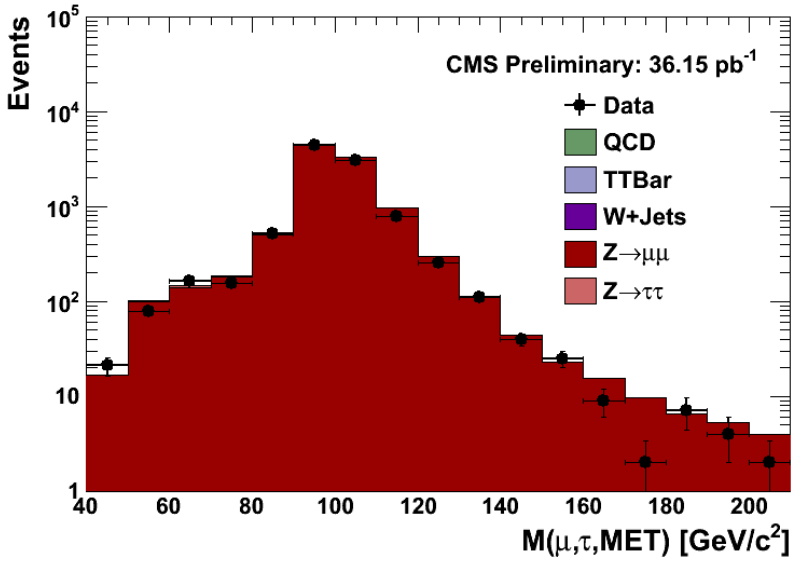

Fig. 17.9. $M(\mu, \tau_h, \not{E}_{\mathrm{T}})$ in the $\mu\tau_h$ $Z \to \mu\mu$ control region.

listed in Table 17.4 and distributions shown in Figure 17.9. For these reasons, it is expected a priori that MC and data will agree. The expected and observed rates in MC and data are consistent. Furthermore, the mass shapes in Figure 17.9 show good agreement. Therefore, not only does the control sample of $Z \to \mu\mu$ serve as a means by which the $Z \to \mu\mu$ contribution in the signal region can be estimated, but also as a validation of the robustness of the analysis and the muon identification criteria.

Table 17.5
$Z \to \mu\mu$ Extraction Efficiencies.

Cut	Data
$\varepsilon^{\not{E}_{\mathrm{T}}}$	0.0029 ± 0.0007
ε^{μ} Veto	0.013 ± 0.0001
Expected Number of Events	0.225 ± 0.054

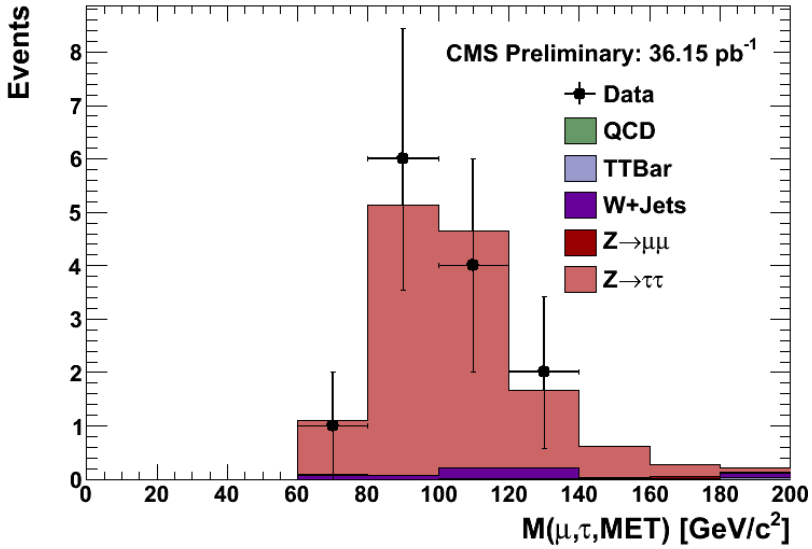

Fig. 17.10. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ in the $\mu\tau_h$ $Z \to \tau\tau$ control region.

17.3 $t\bar{t}$ Background Estimation from Data

As discussed in section 13.8, the main discriminators against W + jet(s) events are the $\Delta\phi(\mu, \tau_h)$ and ζ variables. The same is true for $t\bar{t}$ events where a real light lepton from the semi-leptonic decay of the $W^{\pm}$ and a fake hadronic tau from the hadronic decay of the second $W^{\pm}$ enter the signal region. Thus, it is only natural to obtain a clean sample of $t\bar{t}$ events by removing the $\Delta\phi(\mu, \tau_h)$ and ζ requirements to enhance events with W bosons. Once these requirements have been made, the most significant contribution to the $t\bar{t}$ enhanced region is due to W + jet(s) events. However, unlike W + jet(s), $t\bar{t}$ events contain a larger multiplicity of b-jets from the top quark decays ($t \to bW$). This characteristic can be used to remove a significant contribution from W + jet(s) events. Therefore, the $t\bar{t}$ enhanced control region is defined by making the following modifications to the final signal selection criteria:

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

Figures 17.11 and 17.12 show the μ-τ-$\not\!\!E_{\mathrm{T}}$ invariant mass and $cos\Delta\phi(\mu, \tau_h)$ distributions in this control region. Table 17.6 shows the observed number of events in data and expected number of events in MC. Although it is not required, agreement is observed in both the rates and shapes between the MC simulated samples and collision data.

The number of $t\bar{t}$ events in the signal region can be calculated by measuring the selection efficiencies from events in the $t\bar{t}$ enhanced control sample. Additionally, given a $t\bar{t}$ event with 2 b-jets, standard probability and statistics allows us to determine the probability to tag exactly 2 b-jets, exactly 1 b-jet, and 0 b-jets:

Table 17.6
Events in the $t\bar{t}$ control region for data and MC.

Sample	Events
Data	28
QCD	0.00 ± 1.57
$W + Jets$	1.40 ± 0.31
$t\bar{t}$	26.95 ± 0.36
$Z \to \tau\tau$	1.74 ± 0.23
$Z \to \mu\mu$	0.13 ± 0.06

$$P(2\ b\text{-jets}) = \varepsilon^{b\text{-tag}} * \varepsilon^{b\text{-tag}} \tag{17.6}$$

$$P(1\ b\text{-jet}) = 2 * (1 - \varepsilon^{b\text{-tag}}) * \varepsilon^{b\text{-tag}} \tag{17.7}$$

$$P(0\ b\text{-jets}) = 1 - P(2\ b\text{-jets}) - P(1\ b\text{-jet}) \tag{17.8}$$

where $\varepsilon^{b\text{-tag}}$ is the b-tagging efficiency as measured in [36]. With the above definitions, the number of $t\bar{t}$ events in the signal region can be estimated as follows:

$$N_{t\bar{t}}^{Signal} = N_{t\bar{t}}^{pure} \frac{P^{TCHEM}(0\ b\text{-jets})}{P^{TCHEL}(1\ b\text{-jet}) + P^{TCHEL}(2\ b\text{-jets})} \varepsilon^{cos\Delta\phi(\mu,\tau)} \varepsilon^{\zeta} \tag{17.9}$$

It is noted that eqs. 17.6 and 17.7 are not exact. The exact probability to tag one or two b-jets in the event has additional contributions due to (1) mistakenly identifying a b-jet as the μ or τ_h candidates, (2) mistakenly identifying light quark jets from the W's as b-jets, and (3) additional jets from initial or final state gluon radiation. However, these effects are negligible, especially given the level of accuracy with which

the $t\bar{t}$ signal contribution can be measured. Table 17.7 shows the measurement of the efficiencies from data, and the expected $t\bar{t}$ contribution in the $Z' \to \tau\tau$ signal region obtained using eq. 17.9. The expected $t\bar{t}$ contribution is measured 0.512 ± 0.302. Ensuring that $t\bar{t}$ is a small contribution to the signal region is extremely important because the $M(\mu, \tau_h, \not{E}_{\mathrm{T}})$ mostly sits in the high mass $\tau\tau$ regions, where a new mass resonance could present itself (see Figure 13.32).

Table 17.7
$\mu\tau_h$ $t\bar{t}$ Extraction Efficiencies.

Cut	Data
$P_\zeta - 0.875 P_\zeta^{vis} > -7$	0.464 ± 0.094
$cos\Delta\phi <= -0.95$	0.107 ± 0.058
$\varepsilon^{b-Tagging}$(TCHEM) [36]	0.455 ± 0.016
$\varepsilon^{b-Tagging}$(TCHEL) [36]	0.562 ± 0.020
Probability to tag ≥ 1 $b-jets$ (TCHEL)	80.82 ± 4.30
Probability to tag 0 $b-jets$ (TCHEM)	29.70 ± 2.50
Expected Number of Events	0.512 ± 0.302

17.4 W+jet(s) Background Estimation from Data

The estimation of the W+jet(s) contribution in the signal region is very similar to the $t\bar{t}$ extraction method discussed in section 17.3. In order to measure $\varepsilon^{cos\Delta\phi(\mu,\tau_h)} + \zeta$, the $\Delta\phi(\mu, \tau_h)$ and ζ requirements are dropped to enhance events with W bosons. To remove any contamination from $t\bar{t}$ events, events are required to have exactly 0 jets tagged as b-jets. Additionally, a cut on $M_T(\mu, \not{E}_{\mathrm{T}})$ can be applied to remove contamination from other background sources such as $Z \to \tau\tau$. Therefore, the first W+jet(s) control region, W+jet(s) Region 1, is defined as follows:

- Begin with the selection criteria outlined in section 16
- Remove the requirement on ζ

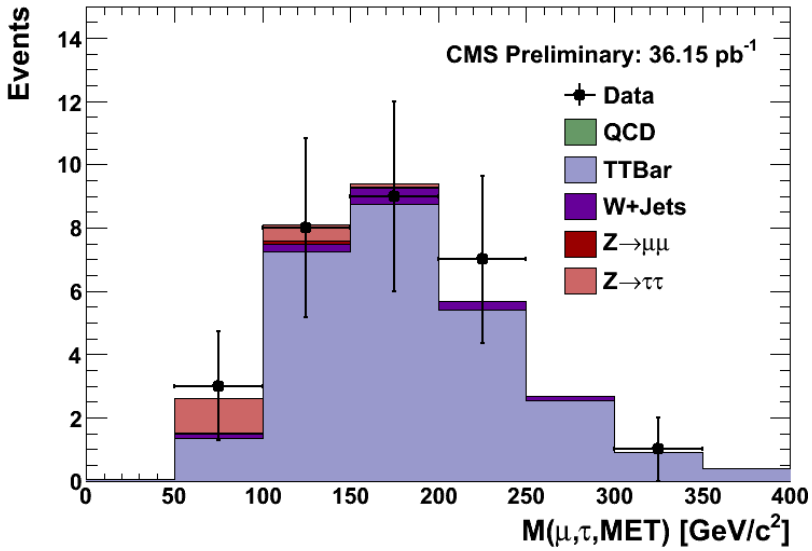

Fig. 17.11. Distributions for $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ in the $\mu\tau_h$ $t\bar{t}$ control region.

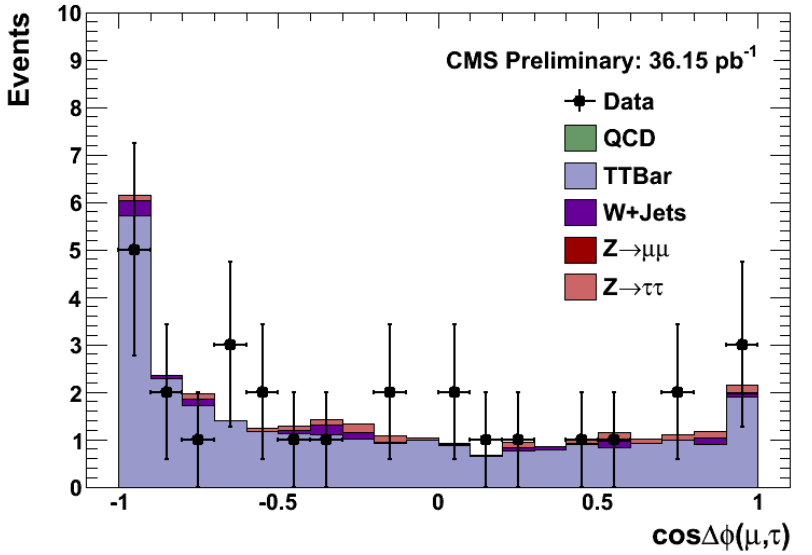

Fig. 17.12. Distributions for $cos\Delta\phi(\mu, \tau_h)$ in the $\mu\tau_h$ $t\bar{t}$ control region.

- Remove the requirement on $cos\Delta\phi(\mu, \tau_h)$
- Define a jet as $\Delta R(\mu/\tau, jet) > 0.5$ and require 0 jets tagged as b-jets using the track counting high efficiency "medium" discriminator (TCHEM)
- $50 < M_T(\mu, \not{E}_{\rm T}) < 100$ GeV

Table 17.8 shows the observed number of events in data and expected events in MC. The W+jet(s) contribution in the signal region can be estimated as follows:

$$N^{Signal}_{W+jet(s)} = N^{pure}_{W+jets(s)} \frac{\varepsilon^{cos\Delta\phi(\mu,\tau)}, \text{ and } \zeta}{\varepsilon^{M_T(\mu,\not{E}_{\rm T})}} \tag{17.10}$$

Table 17.8
Events in the W+jet(s) control region 1 for data and MC.

Sample	Events
Data	70
QCD	0.00 ± 1.57
$W + Jets$	67.2 ± 2.17
$t\bar{t}$	4.08 ± 0.14
$Z \to \tau\tau$	1.68 ± 0.22
$Z \to \mu\mu$	1.21 ± 0.18

Because $M_T(\mu, \not{E}_{\rm T})$ has been utilized to obtain W+jet(s) control region 1, a second control region must be defined that allows for a measurement of $\varepsilon^{M_T(\mu,\not{E}_{\rm T})}$. This can be achieved by noting the jet faking the hadronic tau jet in W + jet(s) events are typically from recoil jets, initial state radiation, and final state radiation. The p_T spectra are soft on average and steeply falling distributions. This characteristic can be used to select a semi-clean sample of W + jet(s) events by looking at *N-1* distributions of the jet spectra. *N-1* distributions refer to the case where one of the final selection criteria is dropped in order to enhance the background contribution

relevant to that particular selection and/or probe the level of accuracy between the MC simulated samples and data for further understanding. The second W+jet(s) control sample, W+jet(s) Region 2, is defined as follows:

- Begin with the selection criteria outlined in section 16
- Remove the τ_h p_T requirement: $p_T > 5$ GeV/c

There is always some concern in using low p_T jets in any analysis due to the high likelihood that MC does not describe data well at low p_T where the modeling of jet fragmentation is most sensitive. However, as is the case throughout the analysis, data-driven methods are employed to remove the dependence on MC to data agreement. Therefore, although any discrepancies between data and MC must be understood, agreement is not necessary for a robust estimation of the background contribution in the signal region. Figures 17.13 and 17.14 show the μ-τ-$\not\!\!E_{\rm T}$ and $cos\Delta\phi(\mu, \tau_h)$ distributions respectively in W+jet(s) control region 1, while Figure 17.15 shows the $M_T(\mu, \not\!\!E_{\rm T})$ distribution in W+jet(s) control region 2. In all cases, agreement between MC and data is observed in both the rates and the shapes of distributions. The purity of W+jets in the background enhanced regions is $> 95\%$. Table 17.9 shows the measurement of the efficiencies from data, and the expected contribution in the $Z' \to \tau\tau$ signal region.

Table 17.9
$\mu\tau_h$ W+Jets Extraction Efficiencies.

Cut	Data
$P_\zeta - 0.875 P_\zeta^{vis} > -7$ AND $cos\Delta\phi <= -0.95$	0.057 ± 0.028
$50 < M_T(\mu, \not\!\!E_{\rm T}) < 100$	0.899 ± 0.034
Expected Number of Events	4.44 ± 2.19

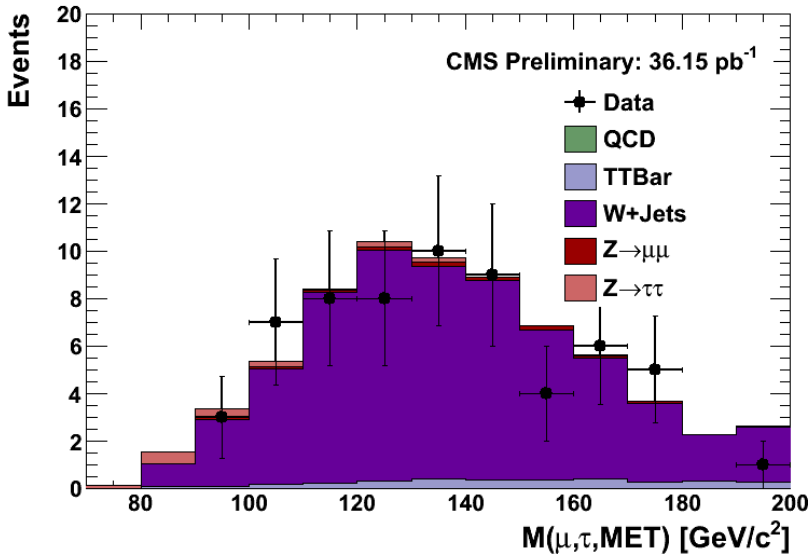

Fig. 17.13. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ in the W+jet(s) Control Region 1.

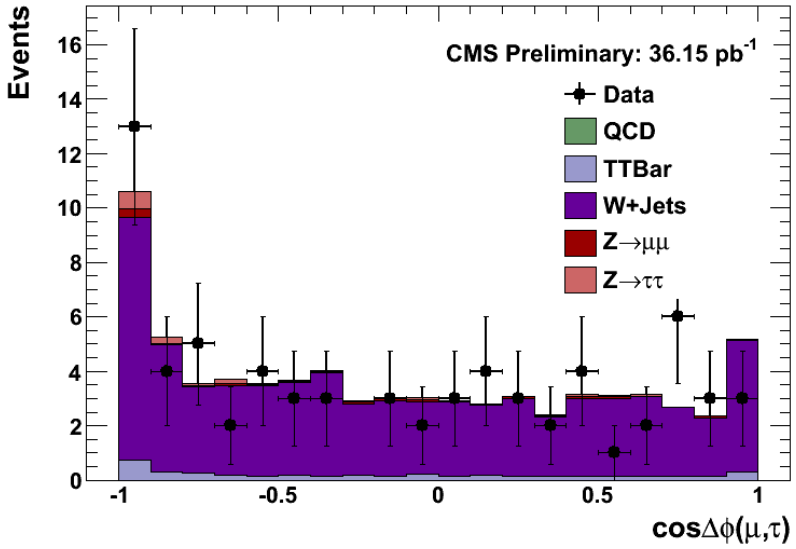

Fig. 17.14. $cos\Delta\phi(\mu, \tau_h)$ in the W+jet(s) Control Region 1.

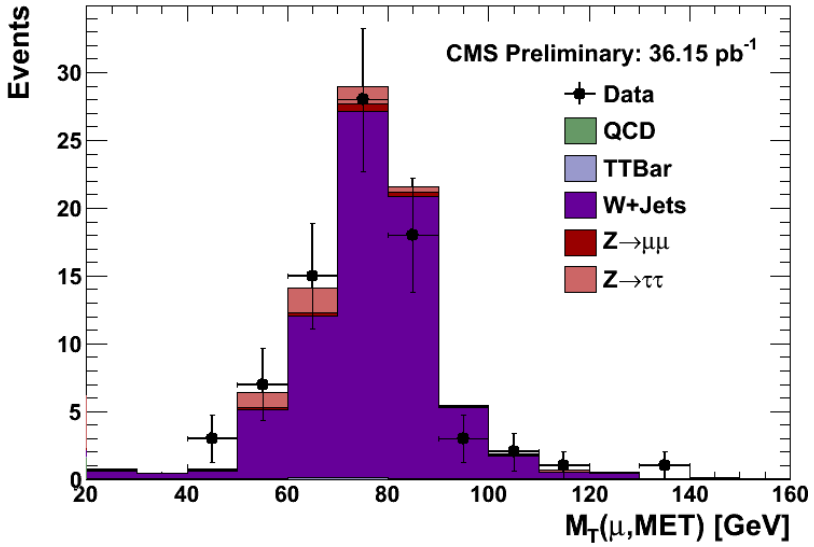

Fig. 17.15. $M_T(\mu, \not{E}_{\mathrm{T}})$ in the W+jet(s) Control Region 2.

17.5 $Z \to \tau\tau$ Validation in Control Region

Obtaining a clean sample of $Z \to \tau\tau$ events is one of the most important aspects of the analysis as it needs to be shown the tau identification criteria employed is successful at identifying taus. Since all the proper corrections and scale factors have been applied to the MC, there is no reason to expect disagreement between data and MC (this is not the case for jet backgrounds). Therefore, obtaining a clean sample of $Z \to \tau\tau$ events where the number of observed events are consistent with the MC expectation ensures the robustness of the analysis and validates the ability to successfully identify taus. A data-driven measurement of the $Z \to \tau\tau$ contribution in the signal region is not employed. Instead, an MC based approach is employed to show that data and MC are consistent in the control region. To validate the MC predictions, the following modifications are made to the final signal selections:

- Modify the requirement on μ p_T: $p_T > 15$ GeV/c
- $M_T(\mu, \not{E}_{\mathrm{T}}) < 40$ GeV

It is important to note that the analysis has been designed to achieve sensitivity in the high mass region, while also enabling us to achieve a relatively clean $Z \to \tau\tau$ signature in the low mass region with only minor modifications to the final selections. Figure 17.10 shows the μ-τ-$\not{E}_{\mathrm{T}}$ invariant mass in the $Z \to \tau\tau$ control region defined above, while the observed events in data and expected events in MC are listed in Table 17.10.

Table 17.10 and Figure 17.10 clearly show that both the $Z \to \tau\tau$ rates and shapes are consistent between data and MC. Therefore, the expected $Z \to \tau\tau$ contribution in the signal region is taken from MC. The expected number of $Z \to \tau\tau$ events in the signal region is 3.98 ± 0.34. Because data and MC shapes/efficiencies are consistent in the control region, the robustness and effectiveness of the tau identification criteria is validated.

Table 17.10
Events in the $Z \to \tau\tau$ control region for data and MC.

Sample	Events
Data	13
QCD	0.00 ± 1.57
$W + Jets$	0.63 ± 0.19
$t\bar{t}$	0.07 ± 0.01
$Z \to \tau\tau$	12.81 ± 0.49
$Z \to \mu\mu$	0.13 ± 0.06

18. DATA IN THE SIGNAL REGION

The signal region consists of events satisfying the criteria outlined in section 16. Table 18.1 summarizes the number of expected and observed events in this signal region. As expected, the largest background source is from W+jet(s) and Drell-Yan production. However, the contamination from these backgrounds is small in the high mass region where a new mass resonance would present itself. The extraction of unbiased background shapes is not possible with only 36.15 pb^{-1} of data. However, as has been shown in the previous sections, suitable cross-checks are made in all control regions to show that the background shapes can be taken from the MC simulated samples. The mass shapes are taken from the simulated samples and fit to obtain smooth trends in the high mass regions. Systematic effects associated with the incomplete knowledge of the shapes are taken in to account in the determination of the final 95% Confidence Level upper limit on the cross-section (see section 19). The signal data distribution (events in data that satisfy all selection criteria), background distributions extracted in a data-driven way, and the signal shape are used to perform a Bayesian fit to extract 95% confidence level upper limits for the possible signal cross-section. Figures 18.1 and 18.2 show the μ-τ-$\not\!\!E_{\mathrm{T}}$ mass distributions in the signal region with Z' masses of 350 and 700 GeV/c^2 respectively. The mass shapes obtained from fitting the MC distributions are normalized to the expected contributions listed in Table 18.1.

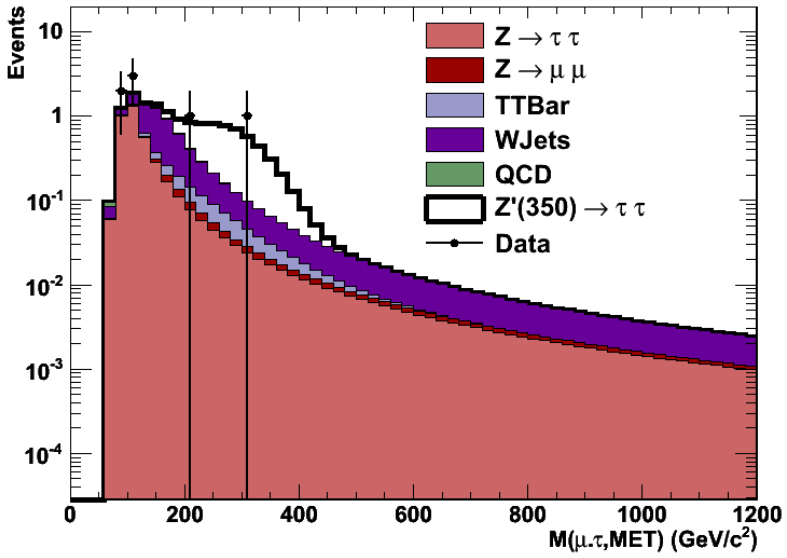

Fig. 18.1. $M(\mu, \tau_h, \not{E}_{\mathrm{T}})$ in the signal region with Z' mass of 350 GeV/c^2.

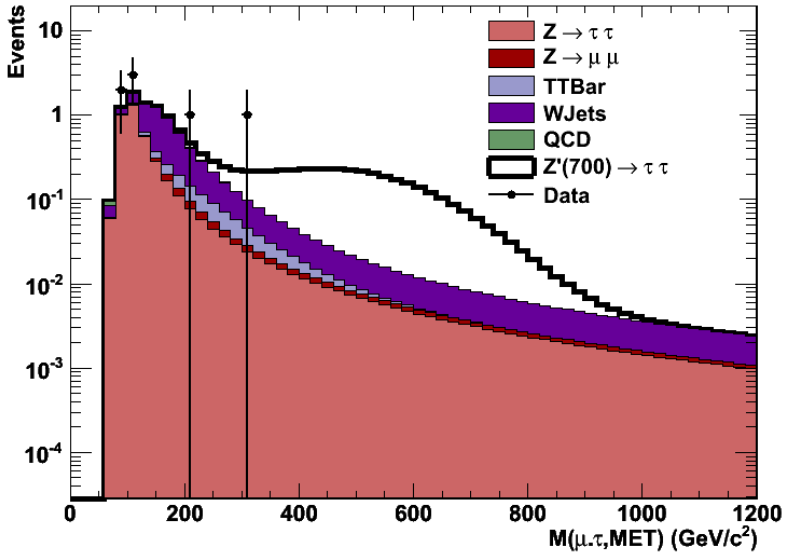

Fig. 18.2. $M(\mu, \tau_h, \not{E}_{\mathrm{T}})$ in the signal region with Z' mass of 700 GeV/c^2.

Table 18.1
$\mu\tau_h$ Events in signal region.

Sample	Events
QCD	0.205 ± 0.101
$W + Jets$	4.44 ± 2.19
$t\bar{t}$	0.512 ± 0.30
$Z \to \tau\tau$	3.978 ± 0.34
$Z \to \mu\mu$	0.225 ± 0.054
Total	9.36 ± 2.985
Observed	7

19. STATISTICAL INTERPRETATION OF THE RESULTS

The statistical interpretation of data and evaluation of signal significance is performed with a Bayesian fit of data against the expected mass distribution to calculate the binned likelihood:

$$\mathcal{L}(\epsilon_1, \epsilon_2, .., \epsilon_n) = \prod_{i=1}^{Nbins} \mathcal{L}_i(\mu_i, \nu_i) \tag{19.1}$$

where $\mathcal{L}_i$ is the Poisson probability of observing events ν_i in data for bin i, given an expectation of $\mu_i(\sigma) = [\text{Background} + \text{Signal}(\sigma)]_i$:

$$\mathcal{L}_i = \frac{\mu_i^{\nu_i} e^{-\mu_i}}{\nu_i!} \tag{19.2}$$

The likelihood distribution is used to obtain the 95% C.L. upper limit for the signal cross section. For sensitivity studies, pseudo-data samples are generated, from background only distributions, using a Poisson based random event generator. The concept of pseudo-experiments and pseudo-data is used to understand the probability for a given outcome to occur. For example, suppose the data-driven background estimation yields an expected total contribution of 9.36. This number only represents the *mean expected* number. This means that if several experiments are conducted, *most* experiments will result in $\sim$ 9 events, while other experiments might result in upper or downward fluctuations. It is likely that some of the experiments will result in e.g. 12 observed events, and some will result in e.g. 6 observed events. In a counting experiment, the probability or *spread* of values is determined by the

Poisson probability density function.

A search that is performed at hadron colliders (like the one presented in this dissertation) represents only one experiment. Because only one experiment is performed, an analysis is designed such that the probability for results to fluctuate are minimized. This is studied by creating *pseudo-experiments* or *pseudo-data* that simulates an ensemble of experiments with each experiment representing a new possible result. Once the ensemble of experiments is created, the spread and mean of the results can be evaluated to gain a quantitative feel for how likely it is that experiment will give rise to a result that unluckily fluctuated upward or luckily fluctuated downward.

Systematic uncertainties can affect the global normalization of the event rate and also create an uncertainty in the knowledge of the mass shape. To incorporate systematic uncertainties into the likelihood calculation, a MC numerical integration method is used to integrate over nuisance parameters. Monte Carlo integration techniques are very common in many applications to obtain an estimate of an integral that is otherwise difficult to evaluate by other means. For example, suppose the evaluation of the following integral is required:

$$\int_a^b f(x)\,\mathrm{d}x. \tag{19.3}$$

The integral can be evaluated by choosing n numbers x_i randomly between [a,b] using some probability density function. For each value x_i, the function f can be evaluated and the n values $f(x_i)$ can be used to approximate the integral as follows:

$$\int_a^b f(x)\,\mathrm{d}x = (b-a) < f(x) > \sim \frac{1}{n}(b-a)\sum_i f(x_i) = (b-a)\bar{f}_n \tag{19.4}$$

where $\bar{f}_n$ is the average of the n possible values of $f(x_i)$. Additionally, the central limit theorem dictates that the sum of a large ensemble of independent values is distributed normally (Gaussian). Therefore, the error in the estimate of the above integral is determined by

$$< f(x) >= \bar{f}_n \pm \frac{\sigma_n}{\sqrt{n}} \tag{19.5}$$

where $\sigma_n = (\bar{f}_n^2 - \bar{f}_n{}^2)^{1/2}$ and $\bar{f}_n^2 = \frac{1}{n}\sum_{i=1}^n f^2(x_i)$. Therefore, the integral converges to the correct value as $n \to \infty$. Because it is not feasible to generate such a large number of random values, the value of n is chosen such that a certain desired level of accuracy is achieved. The generalization to multi-dimensional integrals is achieved by generating n random numbers $(\mathrm{x}_i^1, x_i^2, x_i^3, ...)$.

As stated above, to incorporate systematic uncertainties into the likelihood calculation, the MC numerical integration method described is used to integrate over nuisance parameters. The nuisance parameters are merely a representation of how the n random numbers above are generated. The nuisance parameters, α_k, are generated according to a log normal probability density function for normalizations and Gaussian for mass spectrum uncertainties. If ϵ_n is an efficiency with systematic error $\delta\epsilon$, the likelihood integral becomes:

$$\int \mathcal{L}(\epsilon_1, \epsilon_2, .., \epsilon_n) \mathrm{d}^n\epsilon = N^{-1} \sum_{j=1}^{N} \mathcal{L}(\epsilon_1 + \alpha_1^j \delta\epsilon_1, \epsilon_2 + \alpha_2^j \delta\epsilon_2, .., \epsilon_n + \alpha_n^j \delta\epsilon_n) \qquad (19.6)$$

To incorporate the effects of possible variations in shape, a "morphing" procedure is applied on default, unsmeared mass templates, D_i^{def}, to generate variated templates, D_i^j . Taking into account the smeared templates, the likelihood integral is modified to:

$$N^{-1} \sum_{j=1}^{N} \mathcal{L}(\epsilon_1 + \alpha_1^j \delta\epsilon_1, .., \epsilon_n + {\alpha_n}^j \delta\epsilon_n, D_1^{def} + \alpha_1^j \delta D_1^j, ..., D_n^{def} + \alpha_n^j \delta D_n^j) \qquad (19.7)$$

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. Correlations between systematic uncertainties are considered. For example, to incorporate correlations between systematic effects, the nuisance parameters can be modified as such:

$$\alpha_k = f * \alpha_f + g * \alpha_g \qquad (19.8)$$

where f and g represent the correlated and uncorrelated terms respectively. Figure 19.1 shows the default likelihood (no "smearing" or nuisance parameters incorporated) as well as several superimposed likelihood distributions that represent the

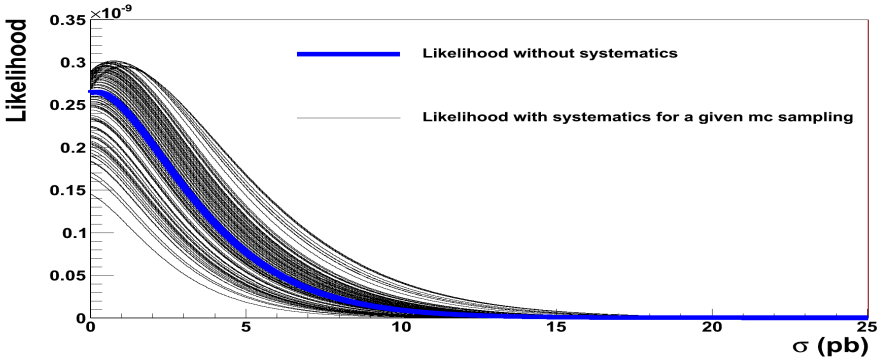

Fig. 19.1. Poisson likelihood for a sample of pseudo-experiments.

affect of the nuisance parameters on the likelihood and the calculation of the 95% C.L. upper limit on the cross-section:

$$0.95 = \frac{\int_0^{\sigma_{95}} L(\sigma)\,\mathrm{d}\sigma}{\int_0^{\infty} L(\sigma)\,\mathrm{d}\sigma}. \tag{19.9}$$

Figure 19.2 shows an example of a MC based sensitivity study where pseudo-data samples are generated from background only distributions as discussed above. Figure 19.3 displays an example pseudo-experiment of the most probable case where the pseudo-data is mostly located in the low mass region, thus giving rise to a 95% C.L. upper limit of $\sigma_{95} = 4.4\ pb$. Figure 19.4 displays an example pseudo-experiment of the "unlucky" scenario where the pseudo-data has fluctuated to the high mass region, thus resulting in $\sigma_{95} = 12.8\ pb$.

The analysis of the search for new heavy mass resonances decaying to τ's described in this dissertation is carried out in the $\mu\tau_h$ final state. However, complimentary studies have also been carried out by the *High p_T Tau Group* at CMS using additional final states: $e\tau_h$, $e\mu$, and $\tau_h\tau_h$. Therefore, although the results of the analysis presented in this dissertation are powerful on their own, the results of this

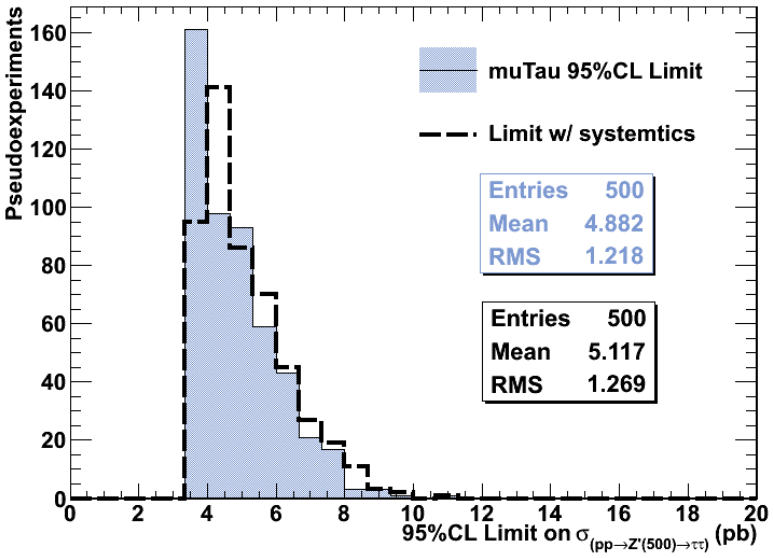

Fig. 19.2. 95% C.L. upper limits for a sample of pseudo-experiments.

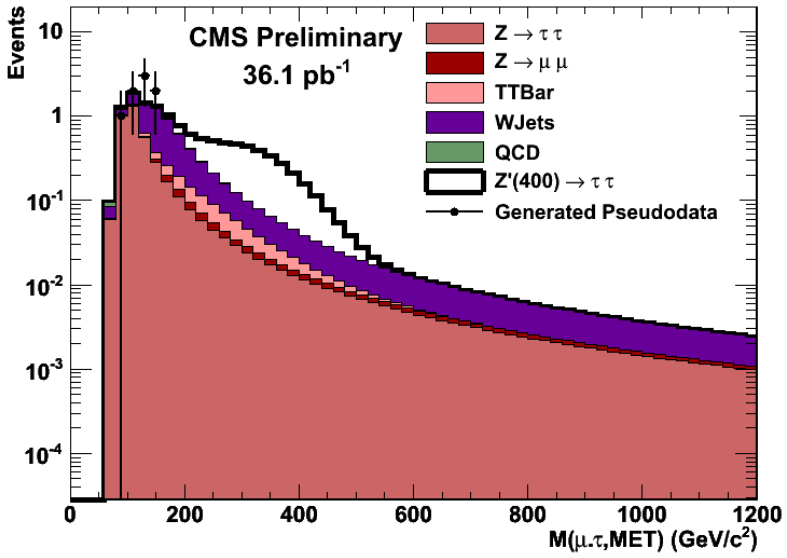

Fig. 19.3. Example pseudo-experiment (with Z' mass of 400 GeV/c^2) resulting in $\sigma_{95} = 4.4$ *pb*.

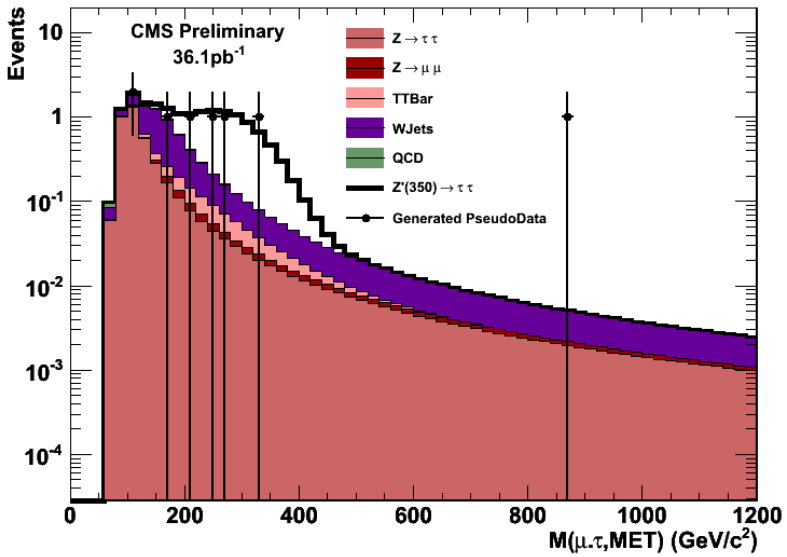

Fig. 19.4. Example pseudo-experiment (with Z' mass of 350 GeV/c^2) resulting in $\sigma_{95} = 12.8\ pb$.

analysis must be combined with the results of the $e\tau_h$, $e\mu$, and $\tau_h\tau_h$ final states to achieve maximum sensitivity [38]. To properly calculate the total/joint likelihood, the likelihoods for individual final states are combined:

$$\mathcal{L}_{total} = \mathcal{L}(\mu\tau_h) * \mathcal{L}(\mathrm{e}\tau_h) * \mathcal{L}(\mathrm{e}\mu) * \mathcal{L}(\tau_\mathrm{h}\tau_\mathrm{h}) \tag{19.10}$$

20. SYSTEMATICS

The main source of systematics is due to the uncertainty in the calculation of the backgrounds (e.g. 33% for $\mu\tau_h$, 45% for $e\tau_h$) and the imprecise knowledge of the luminosity (4%). In cases where a completely data-driven extraction method is employed, the overall systematic uncertainty is mostly driven by the uncertainty in the expected number of events n ($\delta n \sim \sqrt{n}$). In these cases, there is some systematic effect associated with the contamination from other backgrounds in the control regions, however, in most cases, the contamination is negligible. In cases where the contamination is not negligible, the systematic effect due to the contamination is small compared to the systematic uncertainty driven by the lack of high statistics. Systematic effects associated with the incomplete knowledge of the shapes are taken in to account by allowing the fit parameters to fluctuate according to the fit covariance matrix. The W+jet(s), $Z \to \mu\mu$, and $Z \to \tau\tau$ mass shapes are fit with the following functional form:

$$F(x) = \frac{p_0}{p_1 + e^{p_2(x-p_3)} + (x-p_3)^2}, \text{ where } x = M(\tau_1, \tau_2, \not{E}_{\mathrm{T}}) \tag{20.1}$$

The QCD and $t\bar{t}$ mass shapes are fit with the following functional form:

$$F(x) = p_0(p_1 - x)^2 e^{-p_2 x}, \text{ where } x = M(\tau_1, \tau_2, \not{E}_{\mathrm{T}}) \tag{20.2}$$

Figures 20.1 and 20.2 show the mass fits for W+jet(s) and $Z \to \mu\mu$ respectively, while Tables [20.1- 20.4] list the corresponding fit parameters and covariance matrices. Figures 20.3 and 20.4 show the mass fits for QCD and $t\bar{t}$ respectively. Additional

sources of systematics such as trigger efficiencies, identification efficiencies, energy and momentum scale, parton distribution functions, and initial and final state radiation have been included. Systematic effects from muon momentum resolution and scale are mostly driven by the imprecise knowledge of the alignment of the muon system with the rest of the detector and the less accurate determination of muon momentum at high p_T where the magnetic field does not produce highly curved muon trajectories. Systematic effects due to τ_h/jet energy scale and resolution arise to due: (1) uncertainty in the calibration of single particle response used as input to the particle flow algorithm to assign extra photons or neutral charged hadrons when linking tracks to calorimeter clusters; (2) non-linear response of the calorimeter with p_T and η; (3) double counting of energy in cases where the bad matching of tracks to clusters in the particle flow reconstruction algorithm gives rise to additional particles; and (4) effects from the genuine τ constituents (described in the section to follow).

In all cases where the measurements of systematic effects have been carried out by other authors, the POG and CMS recommended values have been used. Scale factors and corresponding uncertainties for the values extracted from simulation are determined using standard tag-and-probe methods [26]. Tau, lepton, and jet energy scale systematics also affect the knowledge of the mass shapes. The $\not\!\!E_{\mathrm{T}}$ scale uncertainties contribute via the jet energy scale and unclustered energy scale. The systematic effects have been determined by "smearing" the values by the corresponding uncertainty and measuring the effect of the newly calculated variables on the event rates and mass shapes. For example, to determine the systematic effect on τ_h p_T, the following equation is used to "smear" the default τ_h p_T value:

$$p_T^{smeared} = k_{scale} \cdot p_T^{generator} + k_{resolution} \cdot (p_T^{reconstructed} - p_T^{generator}) \tag{20.3}$$

where $p_T^{generator}$ is the true generator level value of p_T, $p_T^{reconstructed}$ the default reconstructed transverse momentum, k_{scale} the momentum scale smearing factor, and $k_{resolution}$ the momentum resolution smearing factor. Therefore, to determine the effect of a 3% momentum scale uncertainty on the event yield and mass shape, the τ_h p_T is recalculated with the above equation using $k_{scale} = 1.03$ and $k_{resolution} = 1$. Similarly, to determine the effect of a 5% momentum scale uncertainty on the event yield and mass shape, the τ_h p_T is recalculated using $k_{scale} = 1$ and $k_{resolution} = 1.05$. Notice that setting $k_{scale} = 1$ and $k_{resolution} = 1$ gives the default unsmeared reconstructed value $p_T^{reconstructed}$.

The systematic effect due to imprecise knowledge of the parton distribution functions is determined by comparing CTEQ6.6L PDF with the default PDF and variations within the CTEQ6.6 family of parameterizations. The systematic effect due to imprecise modeling of initial and final state radiation is determined by re-weighting events to account for effects such as missing α terms in the soft-collinear approach [39] and missing NLO terms in the parton shower approach [40].

Table 20.1
Fit parameters for the W+jet(s) mass shape.

Parameter	Fit Result
p_0	$1.46405 \text{ x } 10^3 \pm 1.14757 \text{ x } 10^2$
p_1	$1.66336 \text{ x } 10^3 \pm 1.78424 \text{ x } 10^2$
p_2	$-1.40137 \text{ x } 10^{-1} \pm 7.84007 \text{ x } 10^{-3}$
p_3	$1.47996 \text{ x } 10^2 \pm 2.17566$

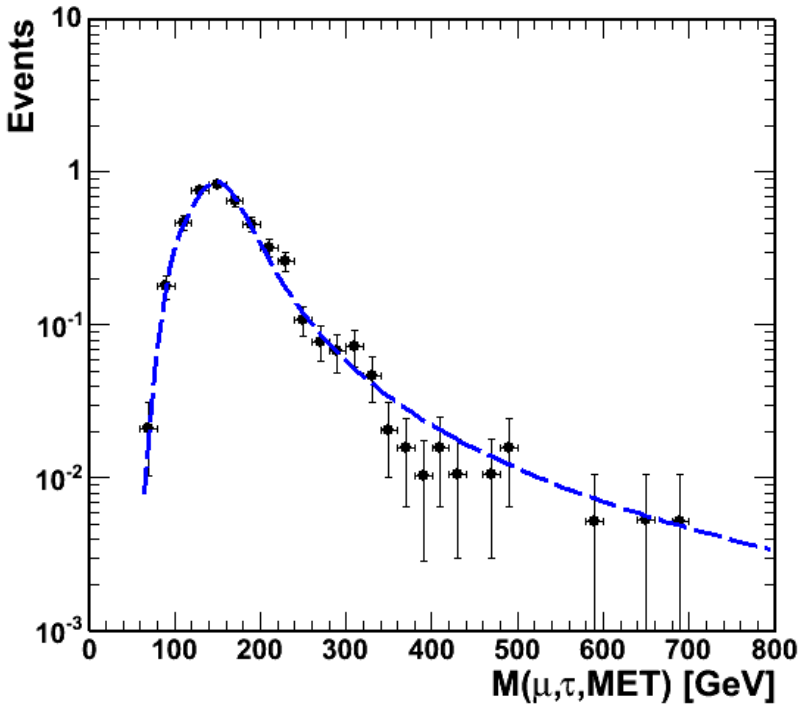

Fig. 20.1. $M(\mu, \tau_h, \not{E}_{\mathrm{T}})$ and corresponding fit for W+jet(s).

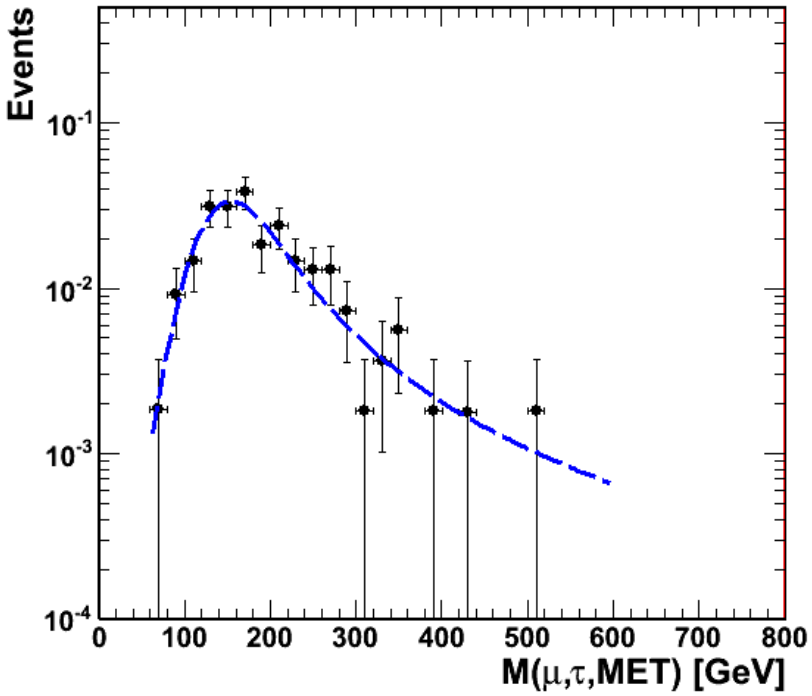

Fig. 20.2. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ and corresponding fit for $Z \to \mu\mu$.

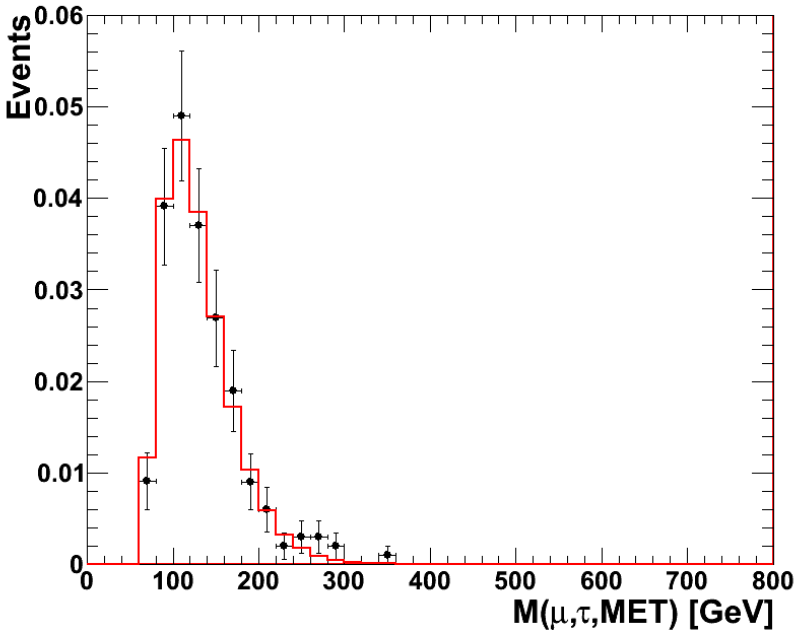

Fig. 20.3. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ and corresponding fit for QCD.

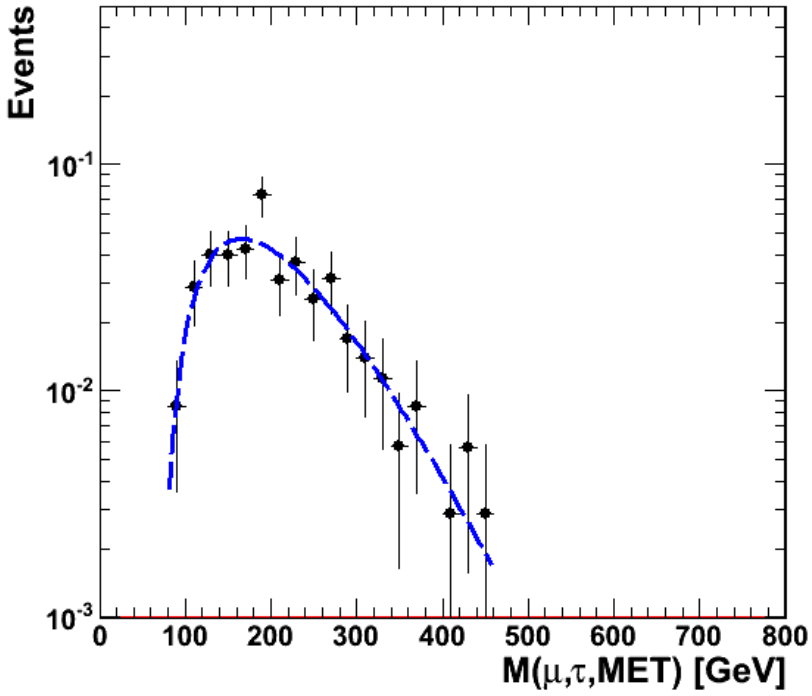

Fig. 20.4. $M(\mu, \tau_h, \not\!\!E_{\mathrm{T}})$ and corresponding fit for $t\bar{t}$.

Table 20.2
Fit parameters for the $Z \to \mu\mu$ mass shape.

Parameter	Fit Result
p_0	$1.15867 \text{ x } 10^2 \pm 2.97788 \text{ x } 10^1$
p_1	$3.52818 \text{ x } 10^3 \pm 1.22067 \text{ x } 10^3$
p_2	$-1.20808 \text{ x } 10^{-1} \pm 1.91469 \text{ x } 10^{-2}$
p_3	$1.60034 \text{ x } 10^2 \pm 7.54570$

Table 20.3
Covariance matrix for the W+jet(s) fit.

	p_0	p_1	p_2	p_3
p_0	$1.32 \text{ x } 10^4$	$1.81 \text{ x } 10^4$	$-3.91 \text{ x } 10^{-1}$	$-7.51 \text{ x } 10^1$
p_1	$1.81 \text{ x } 10^4$	$3.18 \text{ x } 10^4$	$-4.73 \text{ x } 10^{-1}$	$-8.73 \text{ x } 10^1$
p_2	$-3.91 \text{ x } 10^{-1}$	$-4.73 \text{ x } 10^{-1}$	$6.15 \text{ x } 10^{-5}$	$1.28 \text{ x } 10^{-2}$
p_3	$-7.51 \text{ x } 10^1$	$-8.73 \text{ x } 10^1$	$1.28 \text{ x } 10^{-2}$	4.73

Table 20.4
Covariance matrix for the $Z \to \mu\mu$ fit.

	p_0	p_1	p_2	p_3
p_0	$9.06 \text{ x } 10^2$	$3.44 \text{ x } 10^4$	$-2.54 \text{ x } 10^{-1}$	$-6.08 \text{ x } 10^1$
p_1	$3.44 \text{ x } 10^4$	$1.54 \text{ x } 10^6$	-8.16	$-1.77 \text{ x } 10^3$
p_2	$-2.54 \text{ x } 10^{-1}$	-8.16	$3.73 \text{ x } 10^{-4}$	$1.12 \text{ x } 10^{-1}$
p_3	$-6.08 \text{ x } 10^1$	$-1.77 \text{ x } 10^3$	$1.12 \text{ x } 10^{-1}$	$5.77 \text{ x } 10^1$

20.1 Description of Tau ID Systematics

Table 20.5 summarizes the sources of systematics considered and their affect on the acceptance. Tau identification presents one of the most significant systematics of interest for this analysis. In general, tau identification with the shrinking cone algorithm is subject to the following systematic effects (most significant):

- Track finding efficiency per charged hadron

Table 20.5
Systematics for MC and Data.

Source of Systematic	$\mu\tau_h$	$e\tau_h$	$e\mu$	$\tau_h\tau_h$
Luminosity	4%	4%	4%	4%
Muon Trigger	0.55%	—	0.55%	—
Electron Trigger	—	0.39%	—	—
Tau Trigger	—	—	—	3%
Muon ID	0.59%	—	0.59%	—
Electron ID	—	1.37%	—	—
Tau ID	7.0%	7.0%	—	7.0%
Parton Distribution Functions	3.96%	3.96%	3.96%	3.96%
Initial State Radiation	2.14%	2.14%	2.14%	2.14%
Final State Radiation	1.7%	1.7%	1.7%	1.7%
Tau Energy Scale (3%)	2.1%	2.1%	—	2.1%
Electron Energy Scale (1%)	—	1.8%	1.8%	—
Muon Momentum Scale (1%)	1%	—	1%	—
Tau Energy Resolution	Negligible	Negligible	Negligible	Negligible
Electron Energy Resolution	Negligible	Negligible	Negligible	Negligible
Muon Momentum Resolution	Negligible	Negligible	Negligible	Negligible
Background Estimation	33%	45%	14%	28%

- Track finding efficiency convoluted with the probability for a high p_T three pronged tau to produce tracks that are collinear with each other (if the charged hadrons are collinear, the track finding algorithms can reconstruct only one track from more than one charged hadron).

- Probability for charged pions or neutral pions to "leak" out of the signal region and into the isolation region.

- Probability for underlying-event and pile-up tracks and/or photons to fall into the isolation annulus and spoil the isolation requirement.

- Probability for underlying-event and pile-up tracks to fall in to the isolation annulus and spoil the one or three prong requirement.

- Probability for a three prong tau to become a one prong tau.

It is important to note that the tau identification methods employed in this analysis are chosen such that systematic effects are minimized. Because one prong taus are used, the analysis is not subject to systematic effects due to tracks from three prong taus "leaking" out of the signal region. Additionally, due to the one prong requirement, this analysis does not suffer from cases where the charged hadrons from a three pronged tau are collinear and results in e.g. two reconstructed tracks. However, systematic effects where a three pronged tau does produce charged hadrons that are well separated, but where a one prong tau results due to track finding or badly reconstructed tracks is still present. This effect is expected to be small, even for high p_T taus and is measured with the Monte Carlo simulated samples by matching reconstructed taus to generator level taus ($\Delta R < 0.25$) and determining the number of generator level three prong taus that get matched to a reconstructed level one prong tau. It has been found that approximately 0.74% of the simulated taus are actually three prong taus that got identified as one prong taus. The probability for underlying-event and pile-up tracks and photons to fall into the isolation annulus is the same for electrons and muons (as long as similar cones are defined) and can be measured using e.g. standard $Z \rightarrow \mu\mu$ tag and probe methods. The methodology used to measure tau isolation efficiencies is described in section 15. The probability for underlying-event and pile-up tracks and photons to fall into the signal region and spoil the prong requirement can be at most $\sim$ 0.15 / 0.5 * $\delta\epsilon_{iso}$ (size of the signal region with respect to the size of the isolation cone). Additionally, the uncertainty on the track finding efficiency (4 %) for charged pions is determined by measuring the ratio of neutral charm meson decays to two or four charged particles [41]. Therefore, the systematic uncertainty due to tau identification is expected to be $\sim$ 5%. However, to be conservative and to account for some still unknown effects such as out of cone leakage and showering effects in the Ecal (although believed to be small), the tau POG recommended systematic uncertainty of 7% is used. The systematic uncertainty of 7 % is extracted by performing a fit of the tau identification efficiency fixing the

cross section to the CMS $Z \rightarrow ee$ / $\mu\mu$ cross section measured value and using a semi-clean sample of $Z \rightarrow \tau\tau \rightarrow \mu\tau_h$ events. The methodology used to extract the measurement of tau tagging efficiencies is described in section 15. Although the systematic uncertainty due to tau identification was obtained for the HPS algorithm [30] and not the shrinking cone algorithm used in this analysis, the systematic effects are expected to be smaller for the shrinking cone approach for the reasons described above. Therefore, 7% is a conservative estimate. Table 20.5 summarizes the sources of systematics considered and their affect on the acceptance.

21. ANALYSIS RESULTS

Figures 18.1 and 18.2 show the μ-τ-$\not\!\!E_{\mathrm{T}}$ mass distributions in the signal region with Z' masses of 350 and 700 GeV/c^2 respectively. The observed mass spectrum does not reveal any evidence for $Z' \to \tau\tau$ production. Therefore, 95% C.L. upper limits on the cross-section are determined as a function of Z' mass. Figure 21.1 shows the experimental and expected limits on the cross-section as well as the theoretical cross-section for varying Z' masses. The bands on the expected limits represent the 1σ and 2σ deviations obtained using a large sample of pseudo-experiments where the pseudo-data is obtained from background only distributions using a Poisson based random event generator as described in section 19. For completeness, the signal mass distributions for the $\mu\tau_h$, $e\tau_h$, $e\mu$, and $\tau_h\tau_h$ final states are shown in Figure 21.2 [38]. It is noted that although the $e\mu$ final state is cleaner than the e.g. $\mu\tau_h$ and $e\tau_h$ final states, the upper limit is larger due to the smaller branching fraction of e.g. $\tau_h\tau_h$ to $e\mu$. To determine the upper limits on the Z' mass times the branching fraction to $\tau\tau$ pairs, the point at which the experimental limit on the cross-section exceeds the theoretical value is determined. One can see from Figure 21.1 that a Z' with mass less than 468 GeV/c^2 can be excluded at 95% credibility level, exceeding the sensitivity from the Tevatron experiments [42] in 2005.

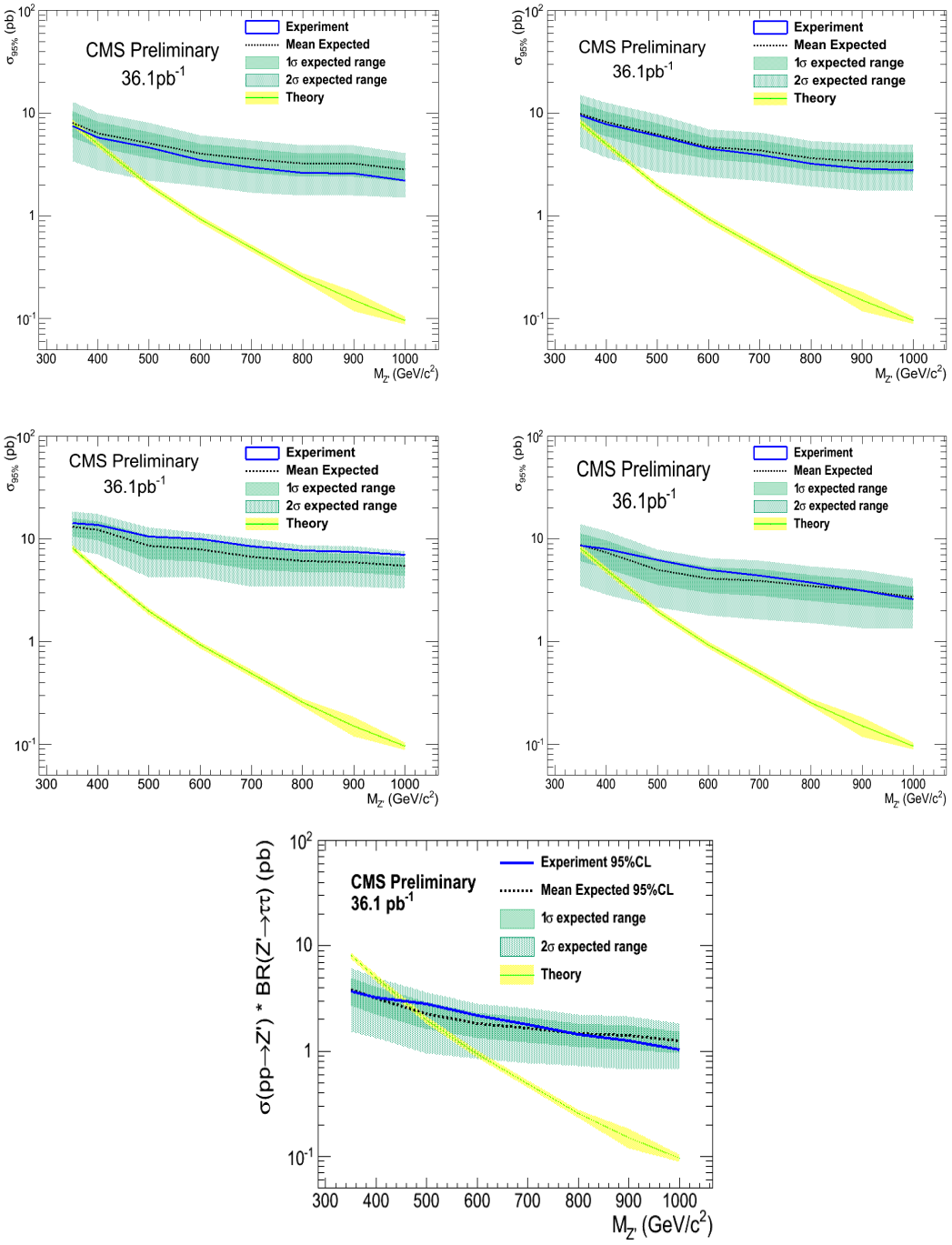

Fig. 21.1. 95% C.L. upper limits on the cross-section times the branching fraction to $\tau\tau$ pairs for (top left) $\mu\tau_h$, (top right) $e\tau_h$, (bottom left) $e\mu$, (bottom right) $\tau_h\tau_h$, and (bottom) joint = $\mu\tau_h + e\tau_h + e\mu + \tau_h\tau_h$ as a function of Z' mass.

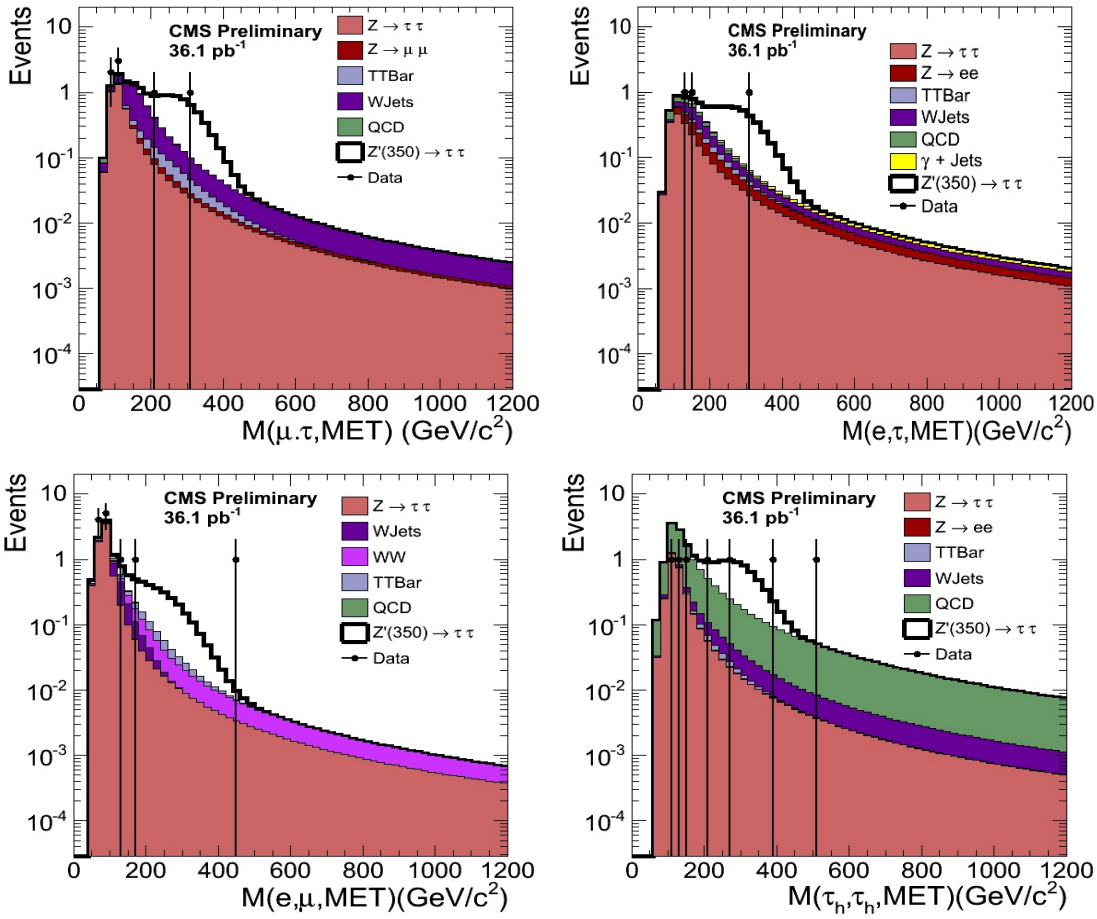

Fig. 21.2. $M(\tau_1, \tau_2, \not\!\!E_{\mathrm{T}})$ distribution for the $\mu\tau_h$ (top left), $e\tau_h$ (top right), $e\mu$ (bottom left), and $\tau_h\tau_h$ (bottom right) final states.

22. CONCLUSIONS

A search for new heavy resonances decaying to $\tau\tau$ pairs using 36.15 pb^{-1} of CMS data from 7 TeV proton-proton collisions has been performed. Monte Carlo simulated samples have been used to understand the SM backgrounds and signal processes and construct a robust analysis that achieves high signal efficiencies but low background rates. Additionally, the optimization of the selections has been performed to maximize sensitivity while minimizing the systematic effects as much as possible. Because the Monte Carlo simulated samples are not expected to agree with observation due to imprecise modeling of effects such as the fragmentation of quarks and gluons, data-driven background extraction methods have been employed to estimate the SM contamination in the signal region. Although agreement is not required in the jet enriched control samples, reasonable agreement has been observed. The efficiencies and distributions for the MC simulated samples have been corrected by applying data-MC scale factors measured using clean samples of muons and taus. Additionally, the MC simulated samples used do not include the simulation of pile-up, therefore, pile-up corrections have been applied to variables such as the transverse momentum imbalance, where pile-up is expected to produce significant discrepancies between data and MC. In cases where agreement is expected, the control samples have been used to validate the robustness of the selections and the muon and tau tagging criteria. W+jet(s) and $Z \to \tau\tau$ are the dominant contributions for the $\mu\tau_h$ final state due to the presence of two real clean τ's for $Z \to \tau\tau$ and a clean muon combined with a hadronic jet that looks like a tau in the case of W+jet(s). Although the expected rates are extracted using data-driven methods, the extraction of the $\tau\tau$ mass shapes is not possible with only 36.15 pb^{-1} of data. Therefore, the shapes are extracted from fits of the MC mass shapes after verifying that data and MC agree in all the necessary control regions. The observed mass spectrum does not reveal any evidence for new heavy resonance production to tau pairs $X \to \tau\tau$. Therefore,

95% C.L. upper limits on the cross-section are determined as a function of Z' mass. The 95% C.L. upper limits are determined using a Bayesian fit of the expected mass shapes against data using a Poisson likelihood. Systematic effects associated to the imprecise knowledge of efficiencies, mass shapes, and additional detector effects have been included in the likelihood calculation via nuissance parameters and integrated out using Monte Carlo integration techniques. These results show that a Z' with Standard Model couplings of mass less than 468 GeV/c^2 can be excluded at a 95% confidence level, surpassing the sensitivity achieved by the CDF Collaboration in 2005 [42].

REFERENCES

[1] WMAP collaboration, *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Determination of cosmological parameters*, Astrophys. J. Suppl. **148** (2003) 175.

[2] M. L. Perl, *Evidence for Anomalous Lepton Production in e^+ - e^- Annihilation*, (1975).

[3] Z. Wan, *Search for high-mass Tau Pairs in 1.96-TeV $p\bar{p}$ collisions*, Rutgers University, 2005 (FERMILAB-THESIS-2005-13).

[4] J. Alcaraz, *ALEPH, DELPHI, L3 and OPAL Collaborations, LEP Electroweak Working Group*, http://lepewwg.web.cern.ch/LEPEWWG/, (2007), 0712.0929, updated for 2008 winter conferences

[5] S. Dawson, *Introduction to electroweak symmetry breaking*, http://arxiv.org/abs/hep-ph/9901280v1, (1999).

[6] P. Langacker, *The physics of Heavy Z-prime Gauge Bosons*, Rev. Mod. Phys **81**, (2009).

[7] A. Breskin, *The CERN large hadron collider: accelerator and experiments*, **Volume 1**, (2008).

[8] Science blogs of physical sciences, *Aerial View of the LHC*, http://www.webelements.com/nexus/aggregator/sources/23?page=4, 2011.

[9] CERN, *CERN Accelerator Complex*, http://atlas.kek.jp/sub/photos/Accelerator/PhotoAccelerator.html, 2011.

[10] T. Ypsilantis, *Particle identification at hadron colliders*, CERN-EP/89-150, (1989).

[11] I. C. Brock, *Electromagnetic shower*, http://www.google.com/imgres?imgurl=http://www-zeus.physik.uni-bonn.de/ brock/feynman/... , 2011.

[12] W. R. Leo. *Techniques for nuclear and particle physics experiments.*, Springer-Verlag, 1994.

[13] R. C. Fernow, *Introduction to experimental particle physics.*, Cambridge University Press, 1986.

[14] W.-M. Yao, *Particle Data Group*, Journal of Physics G **33**, 1 (2006).

[15] R. Arnowitt, B. Dutta, A. Gurrola, T. Kamon, A. Krislock, and D. Toback, *Determining the dark matter relic density in the minimal supergravity stau-neutralino coannhiliation region at the large hadron collider*, Phys. Rev. Lett. **100**, 231802 (2008).

[16] B. Dutta, A. Gurrola, T. Kamon, A. Krislock, A.B. Lahanas, N.E. Mavromatos, and D. Nanopoulos, *Supersymmetry signals of supercritical string cosmology at the large hadron collider*, Phys. Rev. D. (August 2008).

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

[18] V.N. Gribov and L.N. Lipatov, *Deep inelastic e p scattering in perturbation theory.*, Sov. J. Nucl. Phys. **15** (1972) 438.

[19] Z. Was, *TAUOLA the library for tau lepton decay.*, Nucl. Phys. B, Proc. Suppl. 98, 2001.

[20] T. Sj Ostrand, *PYTHIA*, JHEP 2006 026, 2006.

[21] J. Alwall, *MadGraph 5 : going beyond*, http://madgraph.hep.uiuc.edu, 2011.

[22] J. Allison, *Geant4 developments and applications*, http://geant4.cern.ch, 2006.

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

[24] M. Grewal, *Kalman filtering theory and practice using MATLAB*, 2nd ed. New York, John Wiley Sons, Inc., 2001.

[25] W. Adam, *Track reconstruction in the CMS tracker*, CMS Note 2006/041 (2006).

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

[27] CMS collaboration, *Performance of CMS muon measurements in pp collisions at 7 TeV*, CMS Physics Analysis Summary, CMS AN2010/211 , 2010.

[28] R. Ballan, *Muon reconstruction with the CMS tracking system*, Nucl. Phys. B, Proc. Suppl. 177 - 178 (2008) 253 - 254.

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

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

[31] CMS collaboration, *Jet performance in pp collisions at 7 TeV*, CMS Physics Analysis Summary JME-10-003, 2010.

[32] K. Hatakeyama J.P. Chou A. Apresyan, C. Rogan F. Blekman F. Moortgat, K. Theofilatos, J. Weng, M. Pierini, M. Spiropulu W. Andrews, F. Golf, F. Wuerthwein, A. Yagil C. Campagnari, P. Kalavase R. Remington, D. Wang, J. Yelton Ulla Gebbert R. Cavanaugh, F. Lacroix S. Eno, D. Ferencek, F. Santanastasio A. Gurrola, T. Kamon, *Commissioning of uncorrected calorimeter missing transverse energy in zero bias and minimum bias events at* $\sqrt{s} = 900$ *GeV and* $\sqrt{s} = 2360$ *GeV*, CMS Analysis Note AN-2010/029.

[33] CMS Collaboration, *Missing transverse energy performance in minimum-bias and jet events from proton-proton collisions at* $\sqrt{s} = 7$ *TeV*, CMS Physics Analysis Summary CMS-PAS-JME-10-004, http://cms-physics.web.cern.ch/cms-physics/public/JME-10-004-pas.pdf

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

[35] CMS collaboration, *Algorithms for b Jet identification in CMS*, CMS Physics Analysis Summary, http://cms-physics.web.cern.ch/cms-physics/public/BTV-09-001-pas.pdf , 2009.

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

[37] CMS collaboration, *Performance of tau reconstruction algorithms in 2010 data collected with CMS*, CMS Physics Analysis Summary TAU-11-001.

[38] 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, *Search for heavy resonances decaying to tau pairs in CMS*, CMS Analysis Note AN-11-007.

[39] G.Nanava, *How to use SANC to improve the PHOTOS Monte Carlo simulation of bremsstrahlung in leptonic W-Boson decays*, hep-ph/0303260, 2003.

[40] G. Miu, *W production in an improved parton-shower approach*, hep-ph/9812455, 1998.

[41] CMS collaboration, *Measurement of tracking efficiency*, CMS Physics Analysis Summary TRK-10-002, 2010.

[42] D. Acosta, *Search for new physics using high-mass Tau pairs from 1.96 Tev ppbar collisions*, Phys. Rev. Lett. 95, (2005).

VITA

Alfredo Gurrola III was born in El Paso, Texas. He graduated from Socorro High School in June of 2001 and began attending Texas A&M University in August of 2001. He graduated with a Bachelor of Science degree in both mathematics and physics in December of 2005. In the Spring of 2006, he began the doctoral program at Texas A&M University, focusing his work on dark matter searches and searches for new physics with tau leptons at the Large Hadron Collider using the Compact Muon Solenoid Detector. During his doctoral studies, Alfredo Gurrola III conducted research at the Fermi National Accelerator Laboratory in Chicago, Illinois as well as the European Center for Nuclear Research (CERN) near Geneva, Switzerland. He graduated with his Ph.D. in August 2011. His email is: physicsfreddy@gmail.com. His permanent mailing address is: Department of Physics and Astronomy, 4242 TAMU, College Station, TX 77843-4242.