

ABSTRACT

Studies on the Missing Energy Measurement and Supersymmetry Search in 7 TeV Proton-Proton Collisions

Tara A. Scarborough, M.S.

Advisor: Kenichi Hatakeyama, Ph.D.

The main theme of the research presented in this thesis is the search for Supersymmetry, one of the most likely and well-motivated extensions of the standard model of particle physics, in 7 TeV proton-proton collisions delivered by the world's most energetic accelerator — the Large Hadron Collider (LHC). In the Supersymmetry search, the key observable is the “missing transverse energy” (E_T^{miss}), which indicates the signal of the lightest stable supersymmetric particles. These lightest supersymmetric particles are also candidates for the constituents of dark matter observed in astronomical observations. The research presented in this thesis consists of two major parts: (1) improving the measurement of the missing transverse energy, and (2) searching for Supersymmetry with an emphasis on the study of fake E_T^{miss} events and on the interpretation of real E_T^{miss} events using a framework called the “simplified model”.

Studies on the Missing Energy Measurement
and Supersymmetry Search in 7 TeV Proton-Proton Collisions

by

Tara A. Scarborough, M.S.

A Thesis

Approved by the Department of Physics

Gregory A. Benesh, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of
Master of Science

Approved by the Thesis Committee

Kenichi Hatakeyama, Ph.D., Chairperson

Jay R. Dittmann, Ph.D.

Matthew A. Beauregard, Ph.D.

Accepted by the Graduate School
August 2012

J. Larry Lyon, Ph.D., Dean

Copyright © 2012 by Tara A. Scarborough
All rights reserved

TABLE OF CONTENTS

LIST OF FIGURES	vi
LIST OF TABLES	ix
ACKNOWLEDGMENTS	x
DEDICATION	xi
1 Introduction	1
1.1 Standard Model	2
1.1.1 Quantum Chromodynamics	5
1.1.2 Electroweak Interactions	8
1.1.3 Higgs Mechanism and Higgs Boson	8
1.2 Supersymmetry	9
1.3 Jets and Missing Transverse Energy (E_T^{miss})	12
2 Apparatus and Computing	16
2.1 The Large Hadron Collider (LHC)	16
2.1.1 Particle Detectors at the LHC	17
2.2 The Compact Muon Solenoid (CMS) Detector	20
2.2.1 Electromagnetic Calorimeter (ECAL)	21
2.2.2 Hadron Calorimeter (HCAL)	25
2.2.3 Tracker	29
2.2.4 Muon Detector	31

2.3	CMS Computing Model	32
2.3.1	Reconstruction Software and Dataset Format	32
2.3.2	Tiered Computing Structure	33
2.3.3	Dataset Conventions	35
3	ECAL Timing Cuts to Improve Jet and E_T^{miss} Resolution	37
3.1	E_T^{miss} Resolution Studies	39
3.1.1	Tests with the Neutrino Gun Pileup MC Samples	40
3.1.2	Tests with the pp Collision Data	41
3.2	Jet p_T Resolution Studies	48
4	Search for Supersymmetry with Jets and E_T^{miss}	56
4.1	E_T^{miss} Scanning	57
4.1.1	Event Selection for E_T^{miss} Scanning	58
4.1.2	Results from High E_T^{miss} Tail Scanning	60
4.2	Event Selection and Analysis Details	64
4.3	Standard Model Background Contributions	65
4.3.1	$Z(\nu\nu) + \text{Jets}$ Background	66
4.3.2	W and Top Background	67
4.3.3	QCD Background	67
4.3.4	Background Estimation Summary	68
5	Interpretation of the Supersymmetry Search with “Simplified Models”	70
5.1	Calculating Cross Section Upper Limits	71
5.2	Topology Descriptions	74
5.3	Interpretation with Simplified Models	76

5.3.1	Acceptance Times Efficiency for Signal	77
5.3.2	Results	80
6	Conclusion	91
APPENDIX		
A	Additional Material for the Simplified Model Spectra Studies	95
B	Hadronic Calorimeter Reconstruction Software Validation	103
B.1	Software Validation	103
BIBLIOGRAPHY		109

LIST OF FIGURES

1.1	Gluon color composition diagram	7
1.2	Coupling constants of fundamental forces	11
1.3	Diagram of jet evolution and composition	13
1.4	Visualization of cone and k_T jet reconstruction algorithms	13
1.5	Visualization of k_T and anti- k_T jet reconstruction algorithms	15
2.1	CERN accelerator chain and the Large Hadron Collider	18
2.2	Layout of the CMS detector	22
2.3	Layout of the CMS electromagnetic calorimeter	23
2.4	Diagram of an electromagnetic shower	24
2.5	Example of a signal from an energy deposit in the calorimeter	25
2.6	Layout of the CMS hadronic calorimeter and muon chamber	27
2.7	Diagram of a hadronic shower	28
2.8	A drawing of the CMS tracking system.	30
2.9	CMS tiered computing structure flowchart.	34
3.1	Electromagnetic calorimeter barrel and end-cap energy versus time distribution.	39
3.2	Particle flow E_T^{miss} distribution for neutrino gun MC sample	41
3.3	Particle flow E_T^{miss} distributions for the 25 ns early neutrino gun MC sample	42
3.4	Particle flow E_T^{miss} distributions for the 25 ns late neutrino gun MC sample	43

3.5	Particle flow cluster E_T^{miss} before and after implementation of timing cut	45
3.6	Time distribution for EE rechits in high E_T^{miss} events	46
3.7	Event display of high and low E_T^{miss} events	47
3.8	Particle flow E_T^{miss} for L1Mu 25 ns dataset	48
3.9	Particle flow E_T^{miss} for L1EG 25 ns dataset	49
3.10	Jet p_T resolutions for jets in EE for Drell-Yan MC sample with 7 ns cut	51
3.11	Jet p_T resolutions for jets in EB for Drell-Yan MC sample with 7 ns cut	52
3.12	Jet p_T resolutions for jets in EE for photon + jets MC sample with 7 ns cut	53
3.13	Jet p_T resolutions for jets in EE for Drell-Yan MC sample with 10 ns cut	54
3.14	Jet p_T resolutions for jets in EE for photon + jets MC sample with 10 ns cut	55
4.1	Event display of a good “physics” event	59
4.2	Event display showing an ECAL spike	61
4.3	Event display of anomalous events	63
5.1	SMS T1 and T2 topology diagrams	76
5.2	SMS T2tt and T5ZZ topology diagrams	76
5.3	SMS T1bbbb and T1tttt topology diagrams	77
5.4	SMS T1 95% cross section limits, signal efficiency and ISR/FSR uncertainty	81
5.5	SMS T1bbbb 95% cross section limits, signal efficiency and ISR/FSR uncertainty	82
5.6	SMS T2 95% cross section limits, signal efficiency and ISR/FSR uncertainty	83

5.7	SMS T5ZZ 95% cross section limits, signal efficiency and ISR/FSR uncertainty	84
5.8	SMS T1tttt 95% cross section limits, signal efficiency and ISR/FSR uncertainty	85
5.9	SMS T2tt 95% cross section limits, signal efficiency and ISR/FSR uncertainty	86
5.10	SMS T1tttt results comparison between CMS and ATLAS	87
5.11	Best exclusion limits from all CMS searches with SMS	88
A.1	SMS T1 acceptance \times signal efficiency for 14 search bins	96
A.2	SMS T1 signal efficiency uncertainty due to JES, JER, and PDF . . .	97
A.3	SMS T2 signal efficiency uncertainty due to JES, JER, and PDF . . .	98
A.4	SMS T1tttt signal efficiency uncertainty due to JES, JER, and PDF .	99
A.5	SMS T1bbbb signal efficiency uncertainty due to JES, JER, and PDF	100
A.6	SMS T2tt signal efficiency uncertainty due to JES, JER, and PDF . .	101
A.7	SMS T5ZZ signal efficiency uncertainty due to JES, JER, and PDF .	102
B.1	Energy distribution of reconstructed hits for the HCAL end-cap region	105
B.2	Time distribution for the HCAL CaloTowers in the barrel region . . .	106
B.3	HCAL barrel E_T^{miss} distribution and HCAL and ECAL barrel energy distributions	107
B.4	HCAL forward RecHits timing distribution	108

LIST OF TABLES

1.1	List of fermions and gauge bosons in the standard model	3
2.1	Calorimeter absorber and active layer properties	29
3.1	Particle flow E_T^{miss} x projection RMS values	44
3.2	Particle flow E_T^{miss} y projection RMS values	44
4.1	Standard model background contributions to 14 search regions	69
5.1	Summary table of first-generation simplified models	76
5.2	Summary table of third-generation simplified models	77

ACKNOWLEDGMENTS

First and foremost, I would like to thank my mentor, Dr. Kenichi Hatakeyama, for his guidance and advisement. His patience and critical eye has allowed me to become not only a better scientist, but a greater critical thinker in the process. I am proud to have had the opportunity to work with him.

The Baylor CMS group postdoc, Dr. Hongxuan Liu, has been an invaluable resource of information regarding theoretical HEP questions, code debugging, and anything else research-oriented that posed an obstacle to my progress. His help in these matters has been greatly appreciated and his ever-willing nature to assist will never be forgotten.

My committee members Drs. Jay Dittmann and Matthew Beauregard are to be acknowledged as well for their efforts. The pleasant and supportive disposition of both has been a wonderful addition to the madness that (at times) surrounds the process of a defense. I am blessed to have had them agree to be a part of this process with me and help make this document a more complete and well rounded expression of my work.

I would also like to pay homage to my family; they were with me throughout this journey of higher educational pursuits and anything else I could dream up. They are my cornerstone, and I could not have achieved what I have without them and their unyielding support.

DEDICATION

To my matiline

CHAPTER ONE

Introduction

Throughout history, human curiosity has driven innovation in all fields of science. Seeking answers in an attempt to better understand the world around us, we often find that the desired answers create only more questions. The desire to answer What is everything made of? has produced the rich and vibrant research field of elementary particle physics. The field of particle physics developed as a result of this question and through an attempt to break down our current understanding of the macroscopic world into its basic components.

“At quite uncertain times and places,
The atoms left their heavenly path,
And by fortuitous embraces,
Engendered all that being hath.
And though they seem to cling together,
And form “associations” here,
Yet, soon or late, they burst their tether,
And through the depths of space career.”¹

Although he was neither a particle physicist nor a poet, James Maxwell’s “Molecular Evolution” from 1873 provides a stunning image of his understanding of particles and forces written in a language that is intended for a broader audience.

Over the past century, thousands of physicists from around the world have discovered that almost all experimental observations in particle physics can be described by a theory called the standard model. This theory explains hundreds of

¹ Lewis Campbell, *The Life of James Clerk Maxwell*, with a selection from his correspondence and occasional writings and a sketch of his contributions to science (London: Macmillan, 1882): 637–38. QC 16 M4C3 Gerstein Library.

particles and complex interactions with only twelve basic building blocks called fundamental particles, which are controlled by four fundamental forces. Currently the standard model is the best theory describing the fundamental structure of matter and how it interacts.

The theories behind the standard model, as well as a few of its successes and shortcomings, are presented in this chapter.

1.1 Standard Model

The standard model is based around a set of fundamental spin- $\frac{1}{2}$ particles known as fermions that are both point-like and without internal structure. Table 1.1 lists the six leptons and six quarks that are described by this model. The mass of each of these fermions decreases as the list goes down, the heaviest of the constituents being the first listed in each category.² Leptons carry integral electric charges³; the left group having charge -1 and the right having charge 0 . The quarks carry fractional charges; the left group having charge $-\frac{1}{3}$, and the right having charge $+\frac{2}{3}$.

While leptons are seen naturally as free particles (e.g. electrons), quarks strive to be in pairs or a grouping of three. All composite particles are made of quark-antiquark⁴ pairs (mesons) or three quarks (baryons). Mesons and baryons fall into a larger grouping of particles called hadrons, which requires only that the particles are made of quarks and are held together by the strong force.

The most common particles, protons and neutrons, are composed of the lightest u and d quarks in groups of three. Protons consist of uud and neutrons consist of ddu . Heavier particles consisting of s , c , b , t quarks are inherently unstable and

² Possibly except for neutrinos, for which we have been unable to determine their mass.

³ Electric charge comes in multiples of individual small units called the elementary charge, e , approximately equal to 1.602×10^{-19} coulombs.

⁴ An antiquark has the same mass and spin as its quark counterpart, but has differing quantum numbers.

tend to decay rapidly (on the order of 10^{-13} s) to the lighter u and d combinations. The decay of these particles varies dramatically, as the top quark decays before it forms any hadron. The only evidence of these heavier unstable particles on earth is found in cosmic rays and high energy particle collisions.

Table 1.1: List of fermions and gauge bosons in the standard model.

Particles & Forces		
Leptons	Tau (τ)	Tau Neutrino (ν_τ)
	Muon (μ)	Muon Neutrino (ν_μ)
	Electron (e)	Electron Neutrino (ν_e)
Quarks	Bottom (b)	Top (t)
	Strange (s)	Charm (c)
	Down (d)	Up (u)
Gauge Bosons	W Boson (W^\pm)	Z Boson (Z^0)
	Gluon (g)	Photon (γ)

The naming convention for quarks seems a bit incoherent initially, but it follows a certain kind of logic. The ‘up’ and ‘down’ quarks were so named for their isospin component. An isospin vector with $I = \frac{1}{2}$ acts much like a spin vector that has both + ‘up’ or – ‘down’ components. The ‘top’ and ‘bottom’ were found to be the partners of the ‘up’ and ‘down’ quarks, hence their similar naming convention. The name for ‘strange’ quark, however, came from it being a constituent of ‘strange particles’, which were found to be produced primarily through the strong interaction, but decay extremely slowly by way of weak interactions. Lastly, the ‘charm’ quark has no clear reason to be so named, but perhaps was simply a response to its partner named ‘strange’ [1].

Shortly after the existence of quarks was first proposed, another degree of freedom for quarks known as color [2] was suggested in order to explain how quarks could coexist inside some hadrons in otherwise identical quantum states without violating the Pauli exclusion principle. The discussion of quark colors will be deferred to the section in this chapter entitled “Quantum Chromodynamics”.

With the extension of special relativity to quantum mechanics, the Dirac equation predicted that an electron should have a positively charged counterpart. The theory was gradually extended to the idea that every particle has an antiparticle partner with the same mass and spin, but with quantum numbers of opposite sign. For example, the positron is the antiparticle of the electron. It has identical mass, but has a positive charge. If an electron encounters a positron, they annihilate with the transformation of their mass energies into two gamma rays. Going beyond the basics, we can say that an antiparticle is related to the particle by charge conjugation. This includes more than just electric charge; it inverts all internal quantum numbers such as baryon number, lepton number, and strangeness. We denote these particles by either placing a bar on top of the particle symbol, or putting the actual charge with the particle symbol (e.g. for an electron, e , the positron could be signified by e^+ , or for a proton, p , the antiproton is denoted by \bar{p}).

The standard model includes three types of forces acting among particles: the strong, weak, and electromagnetic forces, each working over different ranges and having different strengths. Interactions are often described in terms of the exchange of bosons between the fermions. Each force has its own characteristic force-carrying particle(s): the gluon (strong force), the W^\pm and Z^0 bosons (weak force), and the photon (electromagnetic force).

The strong force acts on any particles that have color (i.e. quarks and gluons), and is necessary for combining protons and neutrons to form nuclei. The massless particle known as a gluon is responsible for mediating interquark interactions of this

type. The weak force underlies some forms of radioactivity, governs the decay of certain subatomic particles, and initiates the nuclear fusion reaction that fuels the sun. The W^\pm and Z^0 bosons mediate interactions of this type. The electromagnetic force works between the electrons and nucleus, but it also holds together neutral atoms via the Van der Waals force. The electromagnetic force is mediated by the photon and is responsible for atomic structure, chemical reactions, the attractive and repulsive forces associated with electrical charge, and magnetism.

Gravity is not yet part of this framework. The quantum theory used to describe the micro world, and the general theory of relativity used to describe the macro world, have proved as yet to be incompatible mathematically [3]. Luckily for particle physics, when it comes to the scale of fundamental particles, the effect of gravity is so weak as to be considered negligible. Only when we have matter in bulk, such as in the macro world, does the effect of gravity dominate. So the standard model still works well for describing known particles and their interactions despite its reluctant exclusion of one of the fundamental forces, which is supposedly mediated by the equally elusive graviton.

1.1.1 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the formal theory of the strong color interaction between quarks and gluons. The new degree of freedom mentioned in the previous section is called “color” and has three possible values: red, blue, or green. Quarks and gluons do not have actual colors, of course, like those encountered in everyday life, but the analogy is based on the idea that all colors are created from three primary colors. Much like antiparticles, color also allows anticolors (cyan: \bar{r} , yellow: \bar{b} , and magenta: \bar{g}) and interactions are assumed to be invariant under color exchange.

The obvious need for such a distinction arose from the definition of the baryon $\Delta^{++} = uuu$ where three identical fermions in a symmetric ground state compose a known particle. This seemed to be a forbidden state by basic Fermi statistics, which led to the conclusion that there must be another missing component to our description. The assumption that all bound states of quarks must be “colorless” implies there is only one combination of colors that will produce a real particle, thus keeping our original description of the standard model intact.

Gluons, which are responsible for mediating interactions between quarks, carry a color and an anticolor. With three colors and three anticolors, we expect $3^2 = 9$ color-anticolor combinations. Specifically, color falls into the special unitary group $SU(3)$ ⁵ (see Figure 1.1), where $3 \otimes \bar{3} = 8 \oplus 1$ results in an $SU(3)$ octet and an $SU(3)$ singlet.

The nine possible color combinations for a gluon are [4]:

$$r\bar{b}, r\bar{g}, b\bar{g}, b\bar{r}, g\bar{r}, g\bar{b}, \frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}), \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g}), \frac{1}{\sqrt{3}}(r\bar{r} + b\bar{b} + g\bar{g}) \quad (1.1)$$

Of the nine different combinations, only eight of these are allowed for a gluon. The first 6 combinations are the basic color-anticolor combinations; the last combination contains each color flavor in equal amounts (resulting in a colorless singlet state). The color-singlet massless gluon would exist as a free particle and give rise to a strong force with infinite range; therefore, it cannot exist. The other two states ($\frac{1}{\sqrt{2}}(r\bar{r} - b\bar{b}), \frac{1}{\sqrt{6}}(r\bar{r} + b\bar{b} - 2g\bar{g})$) are obtained by requiring orthogonality and square root coefficients are introduced where appropriate so that all states are normalized to unity. It is also interesting to note that, unlike the electromagnetic force where the effect of the electric charge decreases as particles are moved further apart, the effect of the color charge actually increases as the distance between particles increases. The potential energy between two such particles increases as they move apart until there

⁵ A set of unitary 3×3 matrices with $\det U = 1$ form the group $SU(3)$.

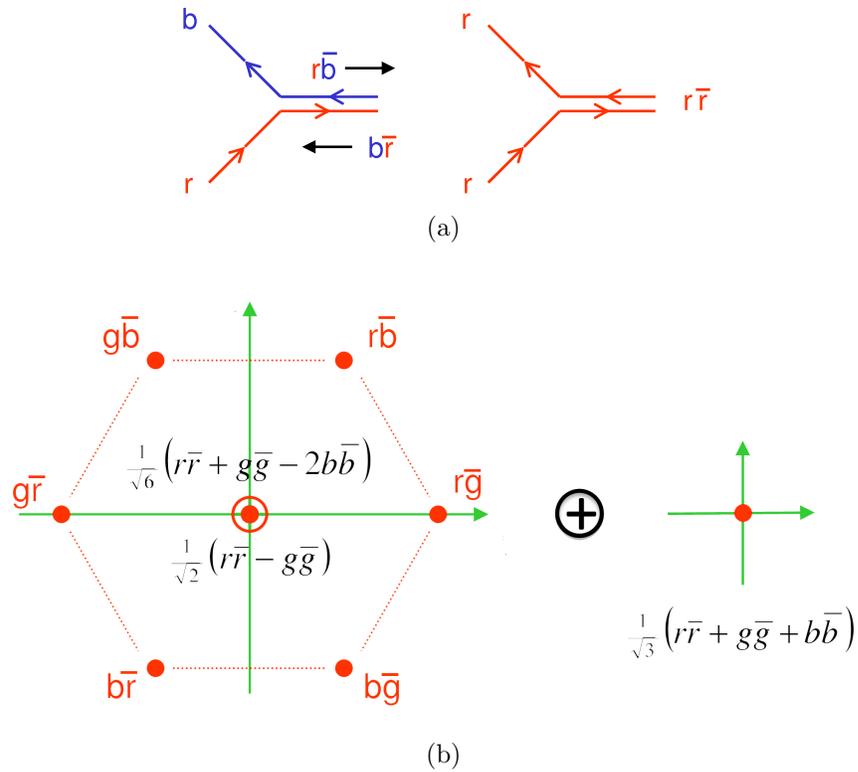


Figure 1.1: (a) Quarks interact via color exchange and color must be conserved at a quark-gluon vertex. The diagrams show two examples of quark interactions, providing a visual explanation of why gluons must contain a color and an anticolor. (b) Color charge falls into the SU(3) group, where the gluon can be represented by 8 of the 9 color combinations seen here (see text for details).

is enough energy to produce a quark-antiquark ($q\bar{q}$) pair. This process is called fragmentation (or hadronization), and can continue as long as there is sufficient potential energy between pairs of colored particles.

1.1.2 *Electroweak Interactions*

In the late 1960's it was discovered that it is possible to describe the electromagnetic and weak interactions as different sides of the same coin. It was predicted that the symmetry between electromagnetic and weak interactions would occur at high energies and the two forces could be described together as a single electroweak force. As the description “electroweak” implies, there is both weak and electromagnetic components to the force. Weak interactions are short-ranged due to the massive mediators of the force (W^\pm, Z^0)⁶ but the electromagnetic force dominates electroweak interactions due to its massless mediator (the photon). This imbalance of power is what induces electroweak symmetry breaking. This particular theory has been extensively tested over the decades with ever-increasing accuracy, solidifying the foundations of a unifying theory for forces [5]. Due to the major success of unifying these two forces, many attempts have been made to unify the remaining forces (Grand Unified Theory — GUT).

1.1.3 *Higgs Mechanism and Higgs Boson*

A common question for particle physicists is ‘What is the origin of the mass of particles?’ The standard model postulates that the Higgs mechanism is responsible for generating particle masses via electroweak symmetry breaking. The broken electroweak symmetry described in the previous section can be explained by the existence of the Higgs boson by the standard model. The Higgs boson is expected to have no spin, electric charge, or color charge and interacts weakly with other

⁶ W^\pm and Z^0 bosons have a very short half-life of $\sim 3 \times 10^{-25}$ s which limits the range of the weak interaction.

particles. The Higgs boson has some limits or constraints set on its mass (theoretical constraints from the electroweak global fit [6], and additional experimental limits from the Large Electron-Positron Collider and Fermilab Tevatron experiment for the last ~ 10 years). Experimental searches for the evidence of the Higgs Boson have been carried out at the Fermilab Tevatron collider and efforts have continued at CERN's Large Hadron Collider with two of its largest detectors: ATLAS and CMS (see Chapter 2). On July 4, 2012, CMS and ATLAS presented findings of a new boson consistent with the Higgs boson. A signal above background predictions was discovered at the mass of 125 GeV, which is consistent with theoretical predictions for this particle. This signal may be another new particle altogether; however, it is widely considered that the Higgs boson discovery would validate the standard model as essentially correct.

Despite the standard model's many successes and effectiveness at describing the phenomena within its domain, it is nevertheless considered to be incomplete. Perhaps it is only a part of a bigger picture that includes new physics that has so far been hidden from our view in the subatomic world. Experiments at the Large Hadron Collider may shed some light on these questions.

1.2 *Supersymmetry*

Physicists attempting to unify the fundamental forces have come to a startling prediction: every fundamental particle should have a massive "shadow" force-carrier particle whose spin differs by $\frac{1}{2}$, and every force carrier should have a massive "shadow" particle. This relationship between fermions and bosons is called Supersymmetry (SUSY). Supersymmetry is a theoretically well motivated extension of the standard model, which may shed light on important questions in particle physics such as the nature of dark matter and the unification of electroweak and strong forces.

The supersymmetric model expects that every normal standard model particle has a supersymmetric counterpart. Quarks, leptons, and neutrino partners are defined as squarks, sleptons, and sneutrinos; and the partner of the gluon and charged (neutral) vector bosons are the gluino, and charginos (neutralinos). One supersymmetric extension of the standard model is defined as the Minimal Supersymmetric Standard Model (MSSM), and is definite on the types of particles that should be present for the theory to be correct, but the broken symmetry implies there is a great deal of uncertainty as to the exact mass scale at which we should see new particles and what experimental signatures they will produce. Supersymmetry has to be a broken symmetry, otherwise the partners of the standard model particles would be of the same mass and evidence of the partners would already have been found.

Force unification is of major importance when attempting to narrow down the correct theories. The left panel of Figure 1.2 presents the inverse coupling constant of the electromagnetic, weak, and strong forces as a function of energy. In attributing a relative strength to the four fundamental forces, it has proved useful to quote the strength in terms of a dimensionless coupling constant, α . It is clear from the plot that at no energy scale do the 3 coupling constants converge. The right hand panel is the same plot as what is shown on the left, but with a supersymmetric contribution. Here it is found that on an energy scale of approximately 10^{16} GeV the three coupling constants converge, helping the development of GUT.

In some of the SUSY models, a quantum number which distinguishes standard model and supersymmetric particles is introduced as R -parity:

$$R = (-1)^{3(B-L)+2S} \quad (1.2)$$

where B is the baryon number, L the lepton number, and S the spin. All standard model particles have $R = 1$, while all superpartners have $R = -1$, so a single SUSY particle cannot decay into just standard model particles if R -parity is conserved. If

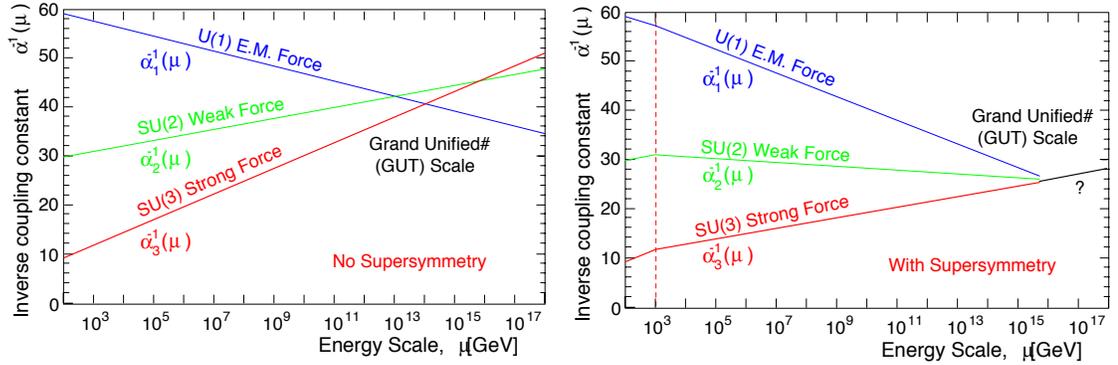


Figure 1.2: (Left) Standard model inverse coupling constants as a function of energy for E.M., weak, and strong forces. (Right) MSSM inverse coupling constants as a function of energy for E.M., weak and strong forces showing unification at high energy [7].

B and L are additively conserved, then it is obvious that R has to be multiplicatively conserved. Due to the lack of concrete information regarding the origin of Supersymmetry, the low-energy Lagrangian necessary to describe the theory contains more than 100 free parameters (e.g. masses, mixing angles, phases, etc.). The addition of R -parity as a constraint helps eliminate many of the “dangerous” terms which would, for example, give rise to a rapid proton decay. Most supersymmetric models assume R -parity conservation which has two important consequences: first that supersymmetric particles must be produced in pairs, and that there must be some lightest supersymmetric particle (LSP). This LSP is usually the lightest neutralino $\tilde{\chi}_1^0$ (see Chapter 4) and is a good dark matter candidate.

Since the LSP can carry away energy without interacting in a detector, the apparent violation of momentum conservation within the detector is an important aspect of SUSY searches. The breaking of the R -parity symmetry would result in lepton and/or baryon number violating processes, and while there are strong experimental constraints on some classes of R -parity violating interactions, others are hardly constrained at all.

At hadron colliders, sparticles can be produced via strong production as, for example, $q\bar{q} \rightarrow \tilde{g}\tilde{g}$ or $gg \rightarrow \tilde{q}\tilde{q}$. Sparticle production is followed by particle cascade decay until a final state is reached containing standard model particles and LSP(s). In each step of the cascade, a standard model particle is produced typically leading to a jet or a lepton in the event. Jets are produced in the cascade decay of the gluino, e.g. $\tilde{g} \rightarrow q\bar{q}\chi_1^0$. Thus, a search for Supersymmetry in events with missing transverse energy (E_T^{miss}) + jets (discussed in the next section) is attractive due to the expected high signal rate. The details of the search will be expanded on in Chapter 4.

1.3 Jets and Missing Transverse Energy (E_T^{miss})

A jet is a collimated stream of particles emitted when a quark (or a gluon) materializes into “stable” particles. We measure the total energy of the particles to try to determine the energy of the original quark, because a knowledge of the latter allows us to make important measurements and new discoveries. Jets are defined by algorithms that cluster together objects such as energies measured in calorimeter towers, particles, or partons (see Figure 1.3). Jet clustering relies on the association of objects based either on proximity in coordinate space (as in cone algorithms) or in momentum space (as in anti- k_T algorithms).

In order to reconstruct jets, two primary methods have been developed: the k_T and cone algorithms (see Figure 1.4). The k_T algorithm has been used successfully at e^+e^- and $e^\pm p$ collider experiments. The cone algorithms have been traditionally used at hadron collider experiments, mainly due to the simplicity in constructing corrections for the underlying event and for multiple interactions in the same bunch crossing.

$$d_{i,B} = k_{T,i}^{2p} \tag{1.3}$$

and

$$d_{i,j} = \min(k_{T,i}^{2p}, k_{T,(i,j)}^{2p}) \Delta R_{ij}^2 / D^2 \quad (1.4)$$

where $k_{T,i}$ is the p_T of the i -th input object, $p = 1$, $\Delta R_{i,j}$ is the distance between each pair of initial objects, and D is the parameter that controls the size of the jet. The symbol B in $d_{i,B}$ in Equation (1.3) refers to the Beam.

If the smallest of these quantities is $k_{T,i}^{2p}$, the corresponding object becomes a jet and is removed from the list of initial objects, and if the smallest quantity is $k_{T,(i,j)}^{2p}$, the two initial objects are merged into a single jet and the original two objects are removed from the list of input objects. This process is iterated until all objects are included in the list of jets [8].

In cone algorithms, objects in a cone in a specified coordinate space are clustered, and the axis of the cone is required to coincide with the direction of the cone defined by a vector sum of all objects inside the cone. Such cones are referred to as stable cones, and jets are formed from these stable cones. Cone algorithms used in experiments so far search for stable cones only from the locations of seeds — objects above a threshold — in order to keep the CPU running time manageable. A list of seed objects is made with the requirement that the p_T of the object exceeds a fixed threshold, which is typically set to 1 GeV.

For the analysis discussed in this thesis, an anti- k_T algorithm is used. The anti- k_T jet clustering algorithm is very similar to the k_T algorithm, but differs where p in Equations 1.3–1.4 corresponds to $p = -1$ (in place of $p = 1$ in the previous case). The primary difference in jets from the anti- k_T and k_T algorithms is that the hard jets are all reconstructed with a circle of radius R with the anti- k_T algorithm, and the softer jets have more complex shapes (see Figure 1.5). Hard jets are those where partons (and/or particles) interact via direct collisions with high momentum transfer, a process which happens rarely when compared to soft interactions. Soft

jets are formed from the partons (and/or particles) that are close enough to interact with one another, but with low momentum transfer.

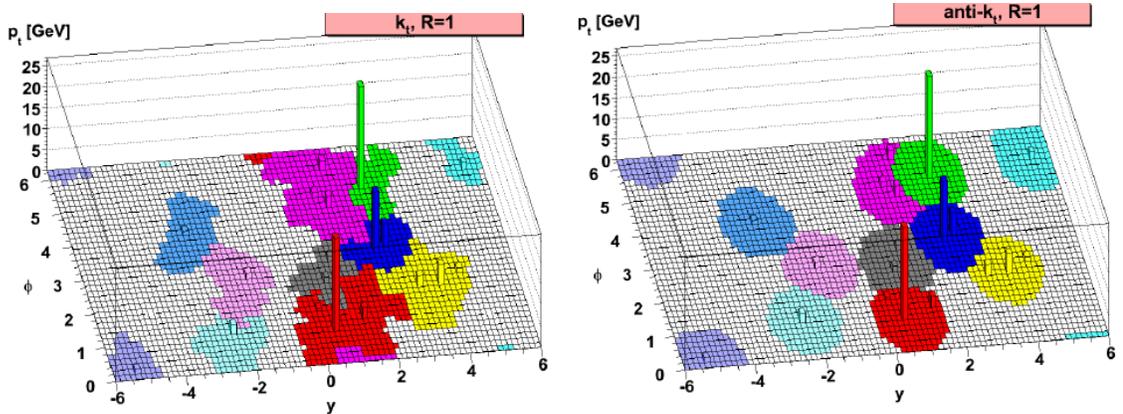


Figure 1.5: A single event showing jet p_T . The left panel reconstructs the jets with the k_T algorithm showing irregular shaped jets. The right panel shows the same jets reconstructed using the anti- k_T algorithm with much cleaner definitions [9].

Missing transverse energy is defined as the magnitude (2-norm) of the vector momentum imbalance in the plane perpendicular to the beam direction, and is denoted as:

$$E_T^{\text{miss}} = |\vec{E}_T^{\text{miss}}|, \quad \vec{E}_T^{\text{miss}} = -\sum_i \vec{p}_{T,i}, \quad (1.5)$$

where the sum is over all the reconstructed particles. E_T^{miss} is a variable used to infer the presence of undetectable particles, such as neutrinos and new weakly interacting, long-lived particles, which makes the calculation of this observable of great importance.

Many beyond-the-standard-model scenarios, including Supersymmetry, predict events with large E_T^{miss} . While E_T^{miss} is an important variable, it is also one of the most complex variables, since E_T^{miss} reconstruction is very sensitive to various detector malfunctions, particles impinging on poorly-instrumented regions of the detector, cosmic-ray particles, and beam-halo particles. All of these factors may result in artificial E_T^{miss} , and are discussed in detail in Chapter 4.

CHAPTER TWO

Apparatus and Computing

2.1 *The Large Hadron Collider (LHC)*

The Large Hadron Collider (LHC) is the world's largest and highest energy particle accelerator and is located at CERN, European Organization for Nuclear Research. CERN, one of the world's largest laboratories for particle physics, is located on the border of Switzerland and France. Founded in 1954, it has hosted many world-class accelerator-based particle experiments of international collaboration.

Protons that are used in the LHC tunnels must undergo a series of accelerations (see Figure 2.1) before they are energetic enough to be collided at the high energies necessary for the current experiments. It all begins with a compressed tank of hydrogen that is connected to the source chamber of the linear accelerator, where electrons are stripped off to leave only the protons. Linac 2 (and Linac 3) are linear accelerators for protons (and lead ions) that provide the initial acceleration of the protons to $1/3$ of c and are fed into the booster stage. The protons are divided into one of the 4 booster rings, accelerated to 91.6% of c , recombined to one group, sent to the Proton Synchrotron (PS), and accelerated to 99.9% of c where a point of transition is reached at which the increased energy cannot be transferred to velocity, where it is converted to an increased *relativistic* mass of the proton,¹ $m\gamma c^2$. At this stage, the energy of each proton has increased to 25 GeV (25 times heavier than a proton at rest). The protons are then channeled into the Super Proton Synchrotron (SPS) which increases the energy of the protons to 450 GeV. Finally,

¹ The kinetic energy of each proton is measured in units of electron volts (eV) and a proton's energy at 99% of c equates to ~ 6 GeV.

they are channeled out into the LHC in one of the two vacuum pipes containing protons which are going in different directions.

Presently the proton collisions occur at the center-of-mass energy of $\sqrt{s} = 8$ TeV, and the energy will nearly double in the coming years. The data used in this thesis was taken during 2011 when the collisions occurred at 7 TeV, but now that the 2012 data taking period has begun, the energy has increased to 8 TeV. These collisions allow us to study the basic constituents of matter — the fundamental particles. It consists of a 27 km ring of superconducting magnets with a number of accelerating structures to boost the energy of the protons along the way. Inside the accelerator, two beams of protons travel close to the speed of light (99.999996% of c) before colliding with one another.

The counter-rotating beams intersect at 4 collision points once they have been further accelerated by the radio frequency (RF) cavities that keeps the proton bunches tightly bunched to ensure high luminosity² at the collision points and hence maximize the number of collisions. The superconducting magnets are used for steering protons in the circular ring. They are built from coils of special electric cable that operates in a superconducting state, efficiently conducting electricity with a minimal resistance or loss of energy.

2.1.1 Particle Detectors at the LHC

There are four locations along the LHC where the protons are collided, and at each location a large-scale detector is housed. At three of these locations, another smaller scale detector shares the collision point. Each detector is described briefly below.

² Luminosity is a measurement of the number of collisions that can be produced in a detector per cm^2 and per second. Larger values of L imply higher number of collisions.

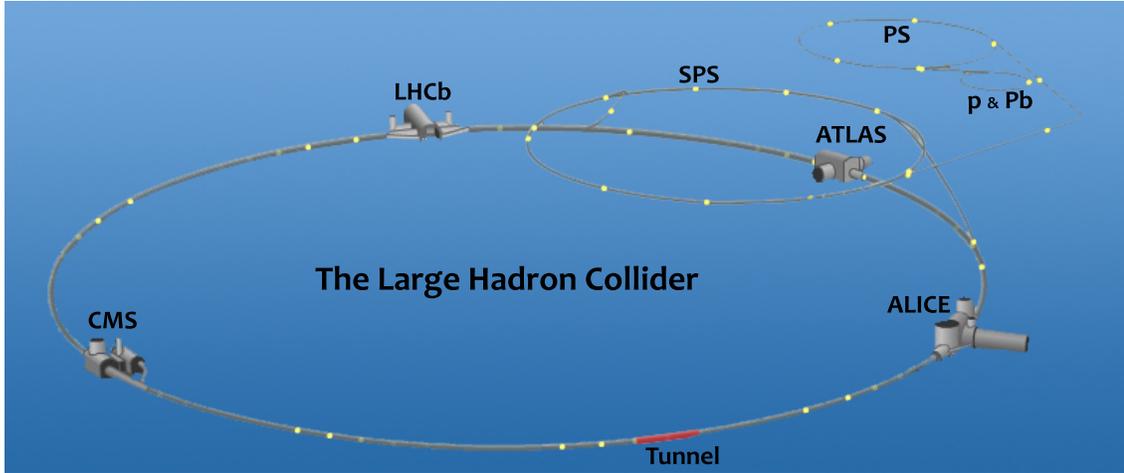


Figure 2.1: Accelerator chain and Large Hadron Collider with the four collision points shown and location of each large scale detector. Linear accelerators Linac 2 and Linac 3 are denoted by p & Pb, the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS) are labeled accordingly, but the Proton Synchrotron Booster is too small to be included here. (Based off of a figure from the BBC World News website [10].)

- CMS (Compact Muon Solenoid) is described in detail in the following section, as it is the main source of the research presented here. This is the primary detector that the Baylor Experimental High Energy Physics group works with and provides contributions to.
- ATLAS (A Toroidal LHC Apparatus) is a large, general purpose detector that is tuned to search for the Higgs boson and signs of new physics. It weighs approximately 7,000 tons, measures approximately 22 m in diameter and 46 m in length, and contains a 2 T solenoid magnetic field for tracking. The calorimeter components include a liquid argon electromagnetic barrel, end-cap, as well as hadronic end-cap and forward calorimeters, scintillating tile barrel and extended barrel. The solenoid is located just in front of the ECAL barrel, which provides a magnetic field critical for the accurate measurement of the p_T of charged particles; however, it also results in significant energy loss by electrons and photons in the material in front of the ECAL. The three superconducting toroid magnets and a set of precise muon chambers, located outside the barrel and end-cap calorimeters, provide a stand-alone muon momentum measurement.
- ALICE (A Large Ion Collider Experiment) is a detector specifically designed to study heavy ion collisions such as Pb-Pb collisions, which are also delivered by the LHC. The purpose of these heavy collisions is to attempt to reproduce processes that occurred shortly ($\sim 100 \mu s$) after the Big-Bang that are contained in a quark-gluon plasma state. Quarks and gluons attract each

other more strongly as they are separated at low temperature and density, which is why no free quarks are found in nature. A quark-gluon plasma is a possible phase of QCD that consists of asymptotically free quarks and gluons. The detector weighs approximately 11,000 tons, measures 8 m in diameter, and 26 m long, and contains a 0.2 T solenoid magnetic field. ALICE contains an ensemble of cylindrical detectors for particle tracking; a silicon pixel detector surrounded by a time projection chamber (designed to cope with the highest conceivable charged particle multiplicities predicted), that is then encased in a transition radiation detector (the main electron detector). For determining the types of particles, a time of flight detector (for intermediate momentum range) is followed by a high momentum particle identification detector. Outside these is a high resolution electromagnetic calorimeter for photon collection and a muon chamber built of multiple resistive plate chambers.

- LHCb (LHC beauty) is a detector designed primarily to study the matter-antimatter asymmetry of the universe. Additionally, its design is also able to perform measurements of B -hadron production cross sections and electroweak physics. The B -hadrons containing b and \bar{b} quarks are unstable and short-lived, decaying rapidly into a range of other particles. Physicists believe that by comparing B and \bar{B} decays, they may be able to gain useful clues as to why nature prefers matter over antimatter. It weighs approximately 6,000 tons, measures ~ 6 m in diameter, and 21 m long, containing a dipole magnet with field strength 4.2 T·m. It consists of a vertex detector, ring imaging Cherenkov detector, and electromagnetic and hadronic calorimeters. What is unique about the LHCb detector is that it is a single arm forward spectrometer with a polar angular coverage from 10 to 300 mrad in the horizontal and 250 mrad in the vertical plane. This provides improved coverage for particles in the forward (particles close to parallel to the proton beams) direction where a large fraction of produced particles head towards.
- TOTEM (Total Cross Section, Elastic Scattering and Diffraction Dissociation) is a small detector that shares the intersection point with CMS. TOTEM aims to obtain precise measurement of the total cross section, elastic scattering and diffractive process. It weighs approximately 21 tons, measures 2.5 m diameter and 440 m in length, and consists of gas electron multiplier (GEM) detectors and cathode strip chambers.
- LHCf (LHC forward) shares an intersection point with ATLAS and consists of two detectors both 30 cm long, 80 cm high, and 10 cm wide. Each detector is located 140 m away from the shared intersection point on either side. The intended purpose of the detector is to measure the number and energy of neutral pions produced by the collider in an attempt to explain ultra-high-energy cosmic rays in collaboration with the Pierre Auger Observatory.

- MoEDAL (Monopole and Exotics Detector At the LHC) is the newest detector added to the LHC in January 2011. It shares an intersection point with LHCb and its primary goal is to directly search for the magnetic monopole and other highly ionizing stable massive particles (or SMPs) at the LHC. It is made up of 400 separate housings that each measure $50\text{ cm} \times 75\text{ cm}$ with a width of 2 cm, weighing in total 3.5 tons. Each housing contains a stack of ten sheets of plastic Nuclear Track Detectors (NTDs) that encase the LHCb vertex detector.

2.2 The Compact Muon Solenoid (CMS) Detector

CMS is a general purpose proton-proton collision detector built on the LHC. The schematic drawing of the CMS detector is shown in Figure 2.2 (a). Approximately, 4,000 people from 180 institutions representing 38 countries form the CMS Collaboration, which operates this detector. It weighs approximately 12,500 tons, measures 15 m in diameter, 20 m in length, and houses a 4 T solenoid magnet.

In particular, the 4 T solenoid magnet sets CMS apart from other competitors as it uses the highest magnetic field with respect to other LHC detectors. It provides a high magnetic field in the tracker volume for precision momentum measurements, and a high enough return flux in the iron outside the magnet to provide a muon trigger and a second muon momentum measurement.

This detector is capable of studying many aspects of proton collisions, and consists of several complementary subsystems in order to measure the energy and momentum of photons, electrons, muons, and many other types of particles produced in the collisions, as shown in Figure 2.2 (b). The main goals of the experiment are to discover the Higgs boson and to look for evidence of physics beyond the standard model, such as Supersymmetry and extra dimensions.

To achieve these goals, the detector was constructed with the following design requirements:

- (1) A high-performance muon system,
- (2) Electromagnetic calorimeters with the best possible energy resolution,

- (3) Hermetic hadron calorimeters,
- (4) High quality central tracking system.

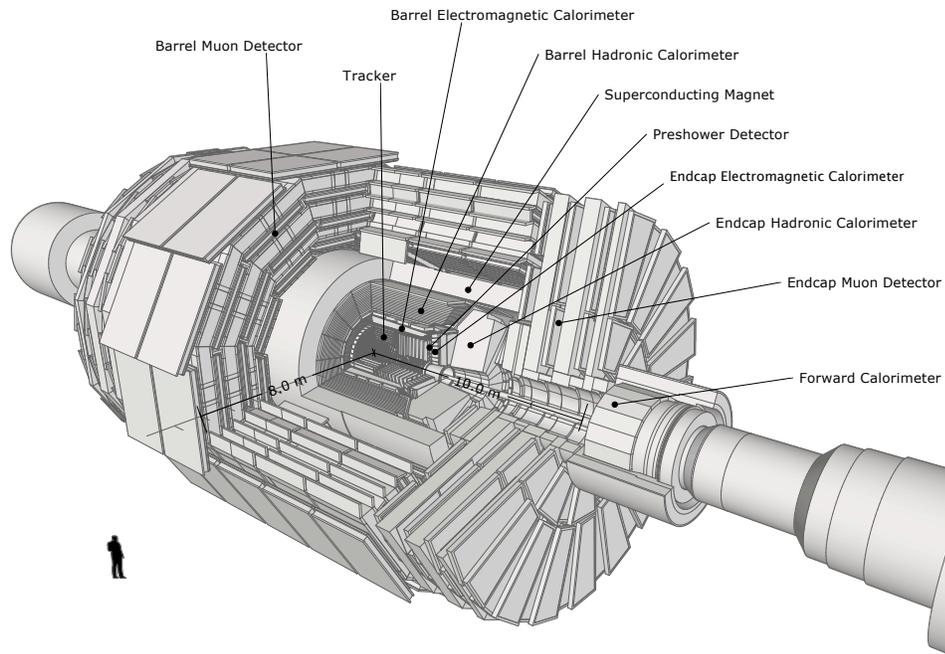
Figure 2.2 (b) shows a transverse slice of the detector from the beam line. Each component is to scale and shows the path of different types of particles that pass through the CMS detector. The components in order of position from the beam line are: the silicon tracker, the electromagnetic calorimeter, hadronic calorimeter, and then the muon chambers. Each will be described in detail in the following sections. The calorimeters played the most significant role in the research presented here and will be described at length first, followed by the tracking components and the muon chambers.

2.2.1 *Electromagnetic Calorimeter (ECAL)*

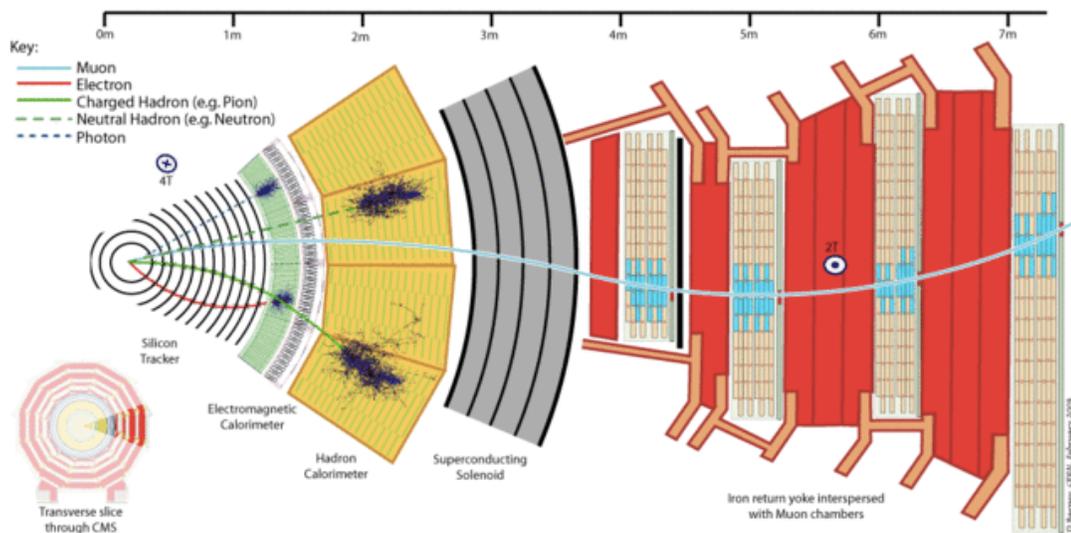
The electromagnetic calorimeter is comprised of a barrel section (EB) and two end-cap regions (EE) and is constructed from approximately 80,000 crystals of lead-tungstate (Figure 2.3). The EB region is built from 5×5 arrays of crystals, known as “supercrystals”, mounted on an aluminum backplate. Each end-cap contains 7,324 lead-tungstate crystals, producing blue-green scintillation light when energy is deposited in the material by charged particles or photons. Each crystal is equipped with a vacuum photodiode tube (VPT) (essentially a single-stage photomultiplier) for light detection.

At the end-caps, the ECAL inner surface is covered by the preshower subdetector, consisting of two layers of lead interleaved with two layers of silicon strip detectors. Its purpose is to increase spacial resolution and aid in pion-photon discrimination.

2.2.1.1. *Electromagnetic Showers and the Calorimeter.* An electromagnetic shower is a cascade of secondary particles produced as a result of high-energy par-



(a)



(b)

Figure 2.2: (a) Overview of the CMS detector to scale. (b) Particle path through the CMS detectors from the pp collision point to the muon chamber (to scale).

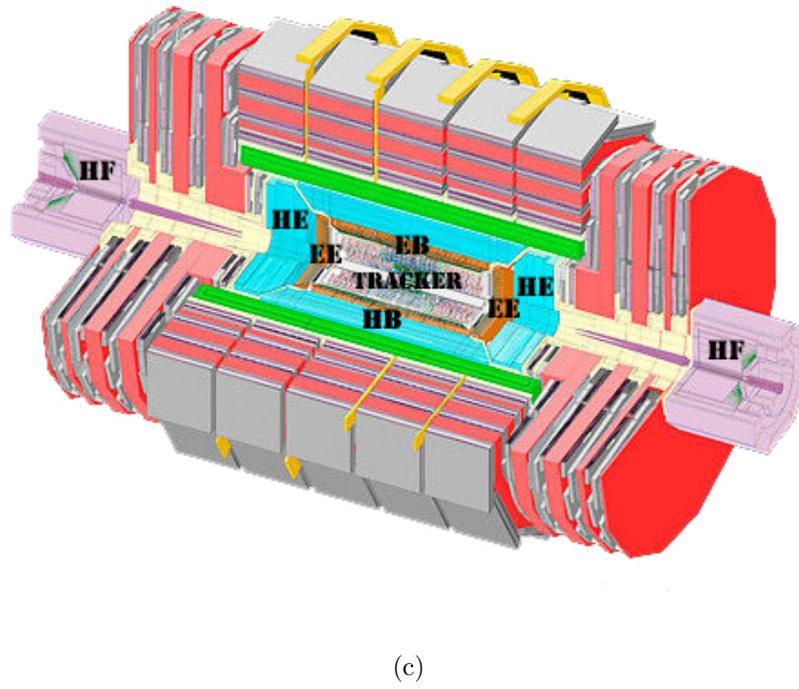
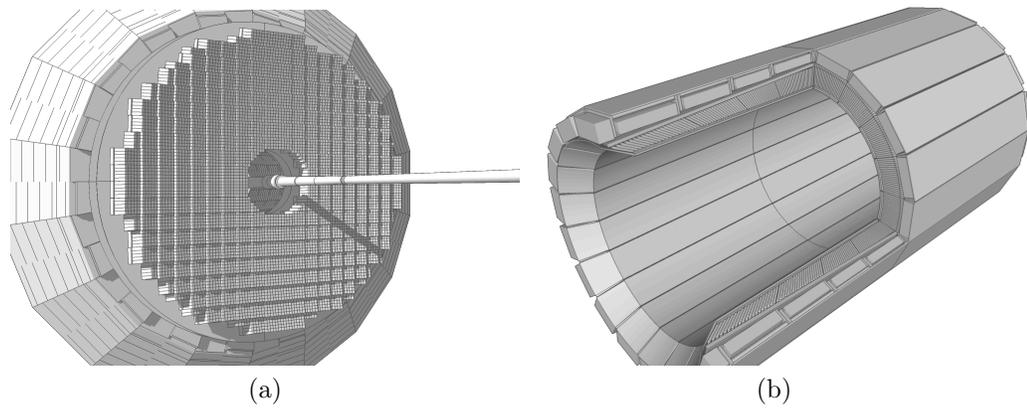


Figure 2.3: (a) EE and HE detector component layout. (b) EB detector component layout. (c) Full CMS detector layout with tracking and calorimeter components labeled.

particles, such as electrons and photons, interacting with dense matter. Incoming particles interact with the dense material and produce multiple new particles with a fractional energy of the incident particle. The process continues until outgoing particles no longer have enough energy to produce secondary particles (see Figure 2.4). Electromagnetic showers are confined to smaller regions in solids that are dense (hence the desirability of using lead) because the higher the atomic numbers, the greater nuclear charges can produce more intense accelerations and so the cascade process can develop more readily than it would otherwise. CMS uses lead-tungstate (PbWO_4) crystals, which have a high density and produce scintillation light in fast, small, well-defined photon showers. This means that the calorimeter system can be very precise and compact in order to fit within the magnetic coil.

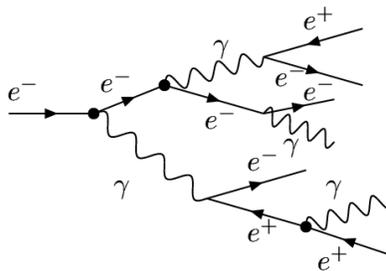


Figure 2.4: Diagram of an electromagnetic shower induced by a single electron.

Electrons lose energy almost entirely through bremsstrahlung, (when energetic particles that pass too close to other nuclei accelerate due to the positive charge of the protons and cause the emission of electromagnetic radiation) where the emitted photons carry off a large fraction of the electron's initial energy. Intense accelerations can produce photons capable of producing more electron-positron pairs (pair production), and secondary particle production continues until photons fall below the pair production energy threshold. Thus, a single electron or photon is the starting point of an avalanche (or electromagnetic shower) of electrons, positrons, and

photons [11]. Electrons and positrons produce electromagnetic showers exactly the same way in a detector.

The left side of Figure 2.5 shows the integrated transverse energy deposition into the calorimeter in $\eta - \phi$ space by an arbitrary particle. The right side is how the same energy deposition would look when distributed among one square supercrystal (each square represents a 5×5 crystal structure called a supercrystal) array of finite size.

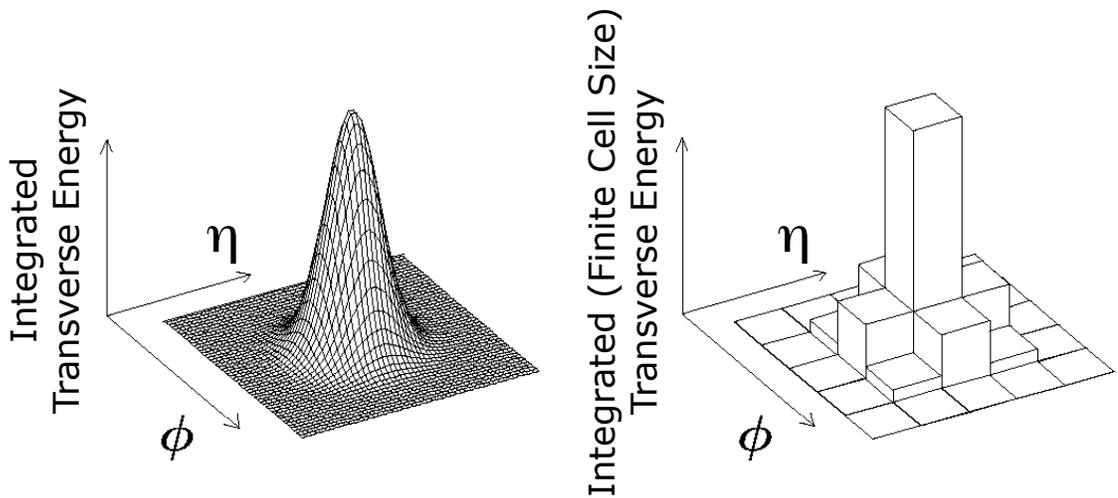


Figure 2.5: (Left) Integrated transverse energy deposition with “infinite” cell size that would occur in ECAL. (Right) Same energy deposition as in the left panel but showing a finite cell size that matches crystal locations and size in $\eta - \phi$ space.

2.2.2 Hadron Calorimeter (HCAL)

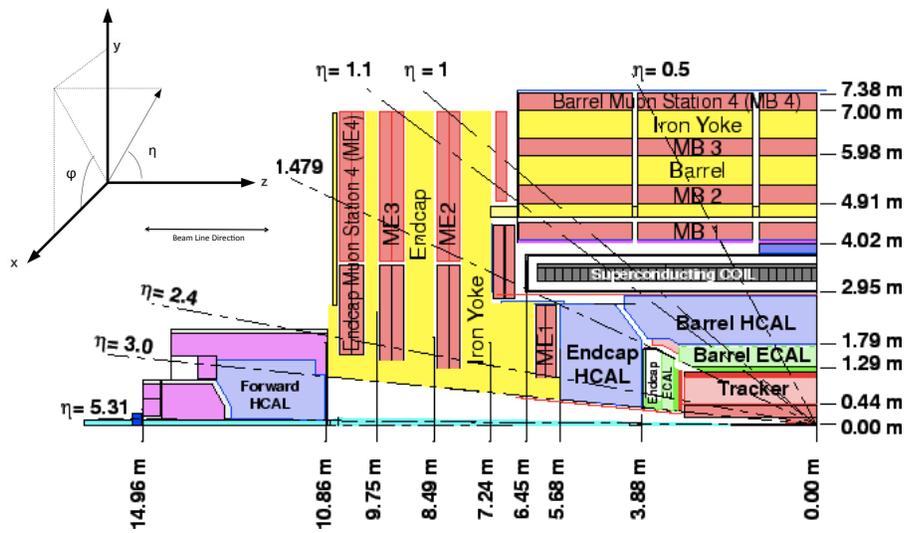
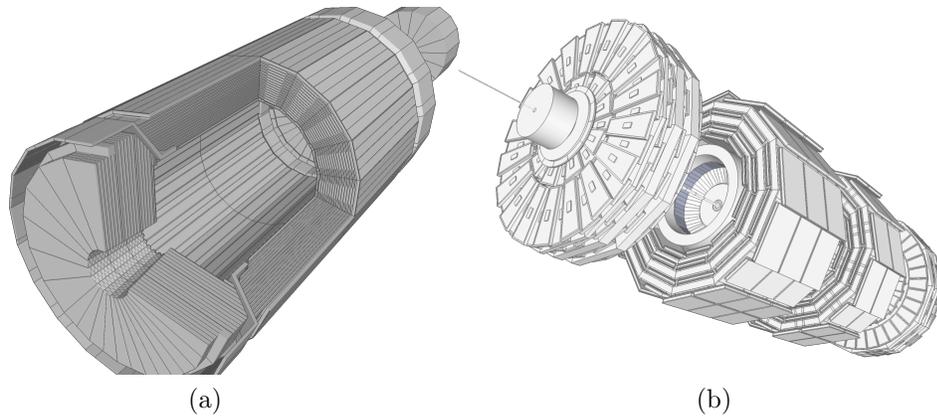
The hadron calorimeter plays an essential role in the measurement of jets, the signature of quarks and gluons, and E_T^{miss} . The E_T^{miss} measurement forms a crucial signature of neutrinos or new particles (such as lightest supersymmetric particles), which are good dark matter candidates. The HCAL finds a particle’s position, energy, and arrival time using alternating layers of absorber and fluorescent scintillator materials that produce a rapid light pulse when the particle passes through, which are then sent to amplified photodetectors and converted to digital signals. The

amount of light in a given region is summed up over many layers of tiles in depth creating an event view of a “tower”. This total amount of light is a measure of an incident particle’s energy.

The hadronic calorimeter is designed to be near hermetic around the interaction region such that events with E_T^{miss} can be identified. It is organized into barrel (HB and HO), end-cap (HE) and forward (HF) sections each composed of 36 wedge structures (see Figure 2.6). The HCAL outer barrel (HO), however, is situated outside of the solenoid magnet coil to measure particles as they leave the magnet and ensures that there is no energy loss that is undetected. The HF receives the bulk of the particle energy contained in the collision so it must be very resistant to radiation and use different materials compared to the other parts of the HCAL. The scintillator and wavelength shifting fibers can degrade due to radiation, thus the HF uses quartz fibers as the active medium to absorb the Cherenkov light³ for signal readouts.

2.2.2.1. *Hadronic Showers and the Hadron Calorimeter.* When hadrons hit atomic nuclei, they lose a lot of energy and this produces a number of secondary particles. Not only does the hadronic shower require a longer time to develop than an electromagnetic shower, but also tends to spread the shower more laterally and longitudinally (see Figure 2.7). Hadrons create signals in the calorimeter when they pass through the absorber materials and the showered particles pass through the active layers. The active layers produce the scintillation light, and when combined with light from all the other layers, it is approximately proportional to the incident hadron’s energy.

³ Cherenkov light is emitted when an electrically charged particle polarizes the medium in which it passes and the molecules emit electromagnetic radiation to return to their ground state.



(c)

Figure 2.6: (a) HCAL detector component schematic. (b) Muon Chamber detector components schematic. (c) Calorimeter and muon chamber component layout in $\eta - z$ space.

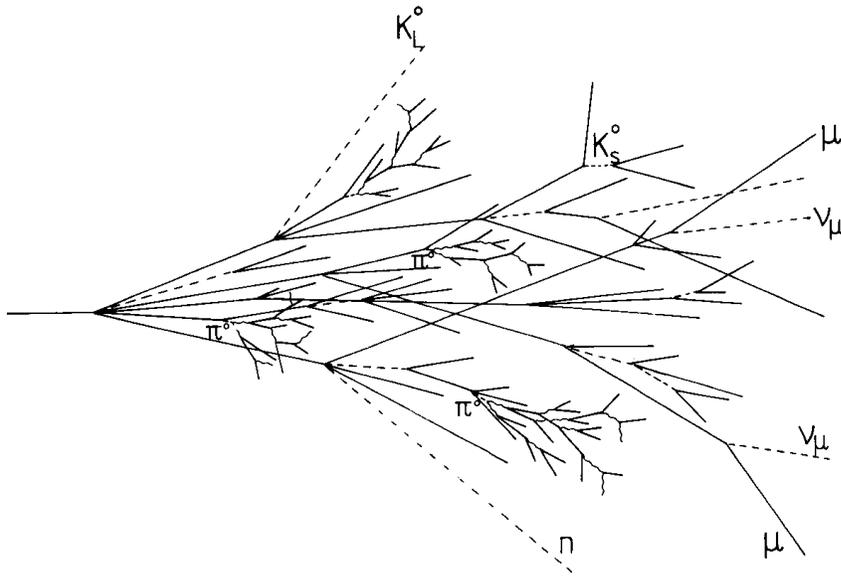


Figure 2.7: Diagram of a hadronic shower in the calorimeter.

Particle showers are measured with fast, low-power instrumentation. The CMS HCAL is a sampling calorimeter, in which there is an interleaving of absorber material and active layers used to periodically sample the hadronic shower. To date, there are no known total absorption detectors that can stop hadronic showers (within $10 \lambda_{\text{int}}$) in a finite thickness (within a few meters or less) and also have a reasonable production cost. Materials such as iron, lead and copper can do this in 1–2 meters, but they are passive materials, which is why we interleave these absorber materials with active layers to form a sampling calorimeter. The CMS HCAL has incorporated brass ($\sim 70\%$ copper) absorbers interleaved with plastic scintillators in the barrel and end-cap regions, and steel ($\sim 98\%$ iron) absorbers with quartz fibers in the forward region. The properties of the absorber and active layers, including their atomic number (Z), density, radiation length (X_0), nuclear interaction length (λ_{int}), and differential MIP energy deposition in a layer of material (dE/dx_{MIP}), are listed in Table 2.1.

Table 2.1: Calorimeter absorber and active layer properties

Material	Z	Density [g/cm ³]	X_0 [cm]	λ_{int} [cm]	dE/dx_{MIP} [MeV/cm]
Fe	26	7.9	1.8	17	11
Cu	29	9.0	1.4	15	13
Plastic Scintillator	–	1.0	42	80	2.0
Quartz	–	2.3	12	43	3.9

Plastic scintillators are a common active layer, where light signals can be brought out for photodetector readout using wavelength shifting fibers. Quartz fibers show reduced radiation damage affecting the region around maximum wavelength PMT sensitivity. This makes them superior to plastic scintillators in the forward region where the components are bombarded with a higher level of damaging radiation.

2.2.3 Tracker

The CMS tracking system represents the world’s largest silicon detector. Silicon detectors provide a significant improvement in momentum and impact parameter resolution for charged particles when compared to other currently-used techniques [12]. Immediately around the interaction point, the inner tracker serves to identify the tracks of individual particles and match them to the vertices from which they originated. The curvature of charged particle tracks in CMS’s strong magnetic field allows their charge and momentum to be measured. Figure 2.8 diagrams how the silicon strips are arranged in 13 layers in the central interaction region, and 14 layers in the end-cap regions, constituting over 76 million channels.

The silicon pixel detector is closest to the beam pipe, meaning in the silicon pixel detector particles emerging from pp collisions tend to be very close to each other. Therefore, it is designed to be able to record millions of particle tracks per

square inch, providing excellent position resolution of incident particles. Outside of this, the silicon strips are aligned along the barrel and end-cap regions. The silicon detectors work in much the same way as the pixels: as a charged particle crosses the material it knocks electron from atoms and within the applied electric field these move giving a very small pulse of current lasting a few nanoseconds. This small amount of charge is then amplified producing a “hit” when a particle passes, allowing us to reconstruct its path.

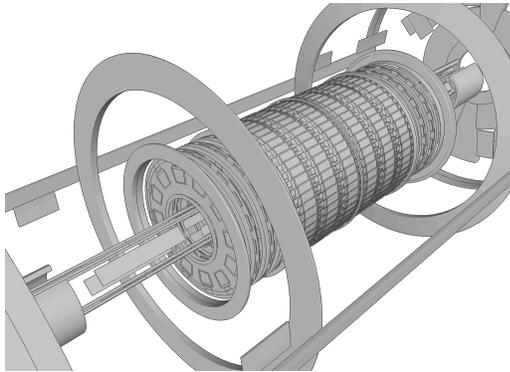


Figure 2.8: A drawing of the CMS tracking system.

2.2.3.1. *Mechanisms of Particle Tracking.* The CMS silicon strip detectors obtain energy measurements from charged particles ionizing electrons in the strips that are collected to retrieve signals from the silicon. They are able to track measurements with a precision of a few μm near the interaction point to improve the momentum measurement and provide a way to determine the decay vertex. The transverse momentum can be determined by:

$$p_{\text{T}} = q \cdot B \cdot \rho \tag{2.1}$$

where q is the particle charge, B the strength of the magnetic field, and ρ the curvature of the particle trajectory. When charged particles enter a magnetic field they will travel in a circular path. The curvature ρ in the above equation is the radius of the circular path. When a charged particle passes through it, the particle

provides enough energy for electrons to be ejected from the silicon atoms, creating electron-hole pairs. Each pixel uses an electric current to collect these charges on the surface as a small electric signal. The curvature can then be used to determine the particle charge and transverse momentum during reconstruction.

2.2.4 *Muon Detector*

Muons are weakly interacting particles that can penetrate several meters of iron without interacting. Thus, in order to detect these particles, a special chamber is required. The CMS muon detector is located outside the solenoid and is interleaved with iron “return yoke” plates intended to absorb any other stray particles that are not muons that may have made it through the calorimeters. There are approximately 1400 muon chambers consisting of 250 drift tubes and 540 cathode strip chambers to track the particle’s position. Specifically, the muon barrel interweaves drift tube chambers (DT) and resistive plate chambers (RPC) and the muon end-caps consist of interweaved cathode strip chambers (CSC) and resistive plate chambers.

2.2.4.1. *Mechanisms of the Muon Detectors.* In the barrel drift tube chamber, when a muon or any charged particle passes through the volume, it knocks electrons off the atoms of the gas (85% Ar, 15% CO₂). These follow the electric field and arrive at the positively-charged wire. The crossing position of the muon can be computed by means of the drift velocity, assuming linear behavior due to a homogeneous electric field. RPCs are fast gaseous detectors that provide a muon trigger system parallel to those of the DTs and CSCs. When a muon passes through the chamber, electrons are knocked out of gas atoms and hit other atoms, causing an avalanche of electrons. The electrodes are transparent to the electrons, which are picked up by external metallic strips. The pattern of hit strips gives a quick measure of the muon momentum. The cathode strip chambers are located in the end-caps where the magnetic field is uneven. The chamber contains positively charged (an-

ode) wires crossed with negatively charged (cathode) strips in a gas volume. When muons pass through the chamber, they kick electrons off the gas atoms that flock to the anode wires creating electron avalanches. The positive ions move towards the cathode strips and induce a charge pulse in the strips at right angles to the wire location, which provide two position coordinates for each passing particle. Each CSC consists of 6 layers, making it able to accurately identify muons and match tracks to those in the inner tracker.

2.3 CMS Computing Model

C++ based computing software (CMSSW) is used to acquire data from this large scale detector. The acquisition of this information seems insurmountable at times when physicists require fine details of particle properties that start out as light being converted to electrical signals. The Worldwide LHC Computing Grid (WLCG) is a global collaboration consisting of more than 140 computing centers in 35 countries, the 4 LHC experiments, and several national and international grid projects. The mission of the WLCG project is to build and maintain a data storage and analysis infrastructure for the entire high energy physics community that uses the Large Hadron Collider at CERN.

2.3.1 Reconstruction Software and Dataset Format

Starting from the most basic electronic output from the detector, information is stored, filtered, and analyzed at various stages to assure accuracy and usability. RAW data (that come directly from the detector) are collected and stored, sent for reconstruction to obtain collision details, and then made available for all users.

The CMS software (CMSSW) used for analysis is constantly updated and many versions have been developed over time. The primary software described in this paper is from the CMSSW_4.2.X series which is designed for the analysis of

data taken during the 2011 run period. CMS consolidates its code base into releases, which are then grouped into release cycles. Each of the releases are dedicated to a specific purpose such as data dating and reconstruction and integration of new *ROOT* [13] or *Fireworks* [14] software versions. The three main data tiers written in CMS are RAW, RECO, and AOD. RAW data contain the full event information from CERN containing “raw” detector information such as element hits. RECO (or “RECO_nstructed data”) is the output from the first pass of processing by the Tier-0. This layer of data contains reconstructed physics objects. RECO can be used for analysis, but the data are too big for frequent or heavy use when CMS has collected a substantial data sample. AOD (or “Analysis Object Data”) is a simplified version of the RECO event information, and is used for most analyses. It provides a trade-off between event size and complexity of the available information to optimize flexibility and speed for analysis.

2.3.2 Tiered Computing Structure

The computing centers available to CMS are distributed in a tiered architecture that functions as a single coherent system. There are 3 tier levels that provide different resources and services to the users. Figure 2.9 provides a visual representation of the essential elements of the flow of real physics data through the hardware tiers.

The first tier (T0) is located only at CERN and performs many functions including accepting RAW data from the CMS Data Acquisition and Trigger⁴ System, repacking this RAW data into primary datasets based on trigger information, and archiving the repackaged RAW data to tape. Afterwards, it acquires calibration constants needed to run reconstruction on the data, feeds the RAW datasets to

⁴ A trigger is a set of requirements that decide which of the millions of collision events are to be kept for analysis. Collection is automatically executed when a particular range of collision types occur during a *pp* collision that are specified by the operator.

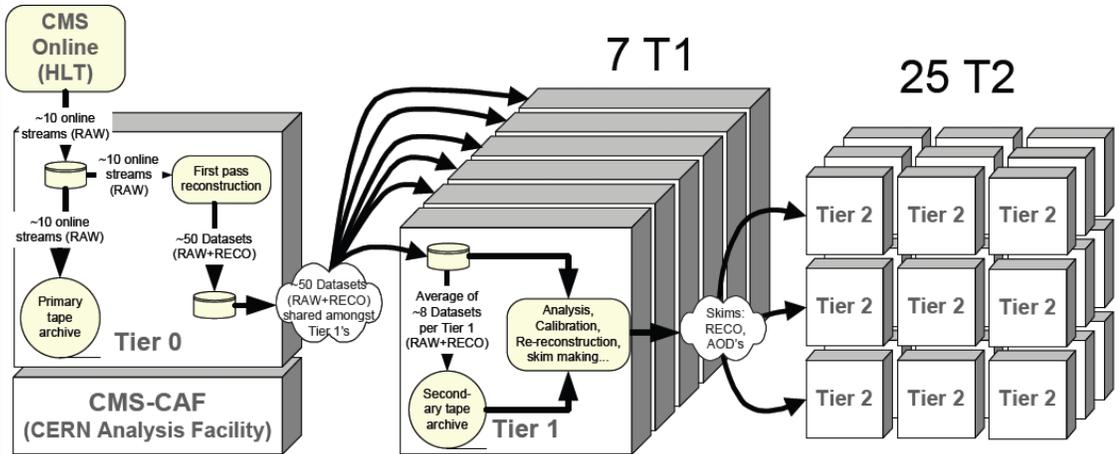


Figure 2.9: CMS tiered computing structure flowchart [15].

the reconstruction. This reconstruction performs a prompt first-pass reconstruction which writes the RECO (and AOD) datasets and then distributes them to T1 facilities.

Seven T1 sites are located in a variety of countries world wide (e.g. USA, Italy, Taiwan, Spain, etc.) to be used for large-scale, centrally organized activities and can provide and receive data from all T2 sites. T1 sites are responsible for receiving a subset of the data from the T0 center, and provide tape archives of part of the RAW (and AOD) data. The data are first divided into Primary Datasets (PDs) that are based on the trigger decision. The primary datasets are structured and placed to make life as easy as possible, e.g. to minimize the need of an average user to run on very large amounts of data. T1 also provides substantial CPU power for scheduled re-reconstruction, skimming, and calibration and distributes these RECOs (and AODs) to T2 centers. Re-reconstruction allows the originally reconstructed events to be re-reconstructed with improved calibrations and/or reconstruction software versions than were used before in order to select a subset of events.

Tier-2 facilities are centered at the university level, but have substantial amounts of CPU resources for user analyses, calibration studies, and Monte Carlo production.

T2 centers provide limited disk space and no tape archiving, but provide services for local communities and allow T2 resources to be available to the whole experiment through the grid. There is a growing number of T2 facilities, and Figure 2.9 is not representative of all current sites.

Tier-3 facilities are smaller, where the local computing structures are available to users for local code development and analysis. The analysis presented in this thesis was performed at Fermilab's T3 as well as a small portion on Baylor University's startup T3 structure. Specifically, the work contained in Chapter 3 was completed solely on Fermilab's T3. Information in Chapter 4 was primarily derived from the Fermilab servers and output was then copied to the Baylor servers (Kodiak⁵) in order to perform the final interpretation for work presented in Chapter 5.⁶

In summary, the Tier-0 center at CERN first reconstructs the full collision events and analysts start to look for patterns. Once CERN has made a primary backup of the data it is then sent to large Tier-1 computer centers in seven locations around the world, where events are reconstructed again using information from the experiment to improve calculations using refined calibration constants. Tier-1 starts to interpret and make sense of the particle events and collate the results to see patterns emerging. Meanwhile each T1 sends the most complex events to a number of Tier-2 facilities (approx. 40) for further specific analysis tasks. Finally T3 locations house user code and specific analysis files to be used locally.

2.3.3 Dataset Conventions

Naming for Data and MC samples follows a certain prescription. As an example, a dataset listed later in this document reads “L1Mu/Run2011B-140ct2011-v1/RECO”.

⁵ <http://www.baylor.edu/lariat/news.php?action=story&story=47126> at kodiak.baylor.edu

⁶ Although Baylor's T3 is not entirely completed and setup for CMS analysis work, initial testing was done and limited calculations discussed in Chapter 5 were performed there.

The first part of the name, `L1Mu`, describes the primary dataset (PD) name. The second part, `Run2011B-14Oct2011-v1` contains several pieces of information regarding the time frame of the data and the reconstruction of the data. `Run2011A` or `Run2011B` states that the data was taken from the 2011 run period and the `A` or `B` signifies whether it was taken earlier or later in the 2011 period. The date indicated, `14Oct2011`, informs the users what date the re-reconstruction was performed to improve the sample over the prompt reconstruction data. This allows the users to check any specifics involved with the particular re-reconstruction code because several different versions may be developed during a specific run range and time frame. Finally, the last part of the name, `RECO` specifies the tier where the data was taken from.

Monte Carlo samples are defined in much the same way. Later we will see the sample:

```
/DYToMuMu_M-20_TuneZ2_7TeV-pythia6/Summer11-HCal25_2TS_PU_S4_START42_V11-v2,
```

where `DYToMuMu` is the Drell-Yan process for $Z \rightarrow \mu\mu$ with a standard model production mass of ~ 20 GeV (`M-20`). The naming convention also includes a tuned `PYTHIA6`⁷ parameter for MC samples of 7 TeV collisions (`TuneZ2_7TeV-pythia6`) that is followed by the time frame of the production (`Summer11`). The remainder of the MC sample name gives specifics about the simulated sample. For this particular sample, the Hadronic Calorimeter, `HCal`, contains two 25 ns time slices, `25_2TS`, with pileup (explained in Chapter 3) simulation described with a specific “Global-Tag” variable of `START42_V11`.

⁷ A general purpose event generator, containing theory and models for a number of aspects of physics, including hard and soft interactions, parton distributions, initial- and final-state parton showers, multiple interactions, fragmentation, and decay.

CHAPTER THREE

ECAL Timing Cuts to Improve Jet and E_T^{miss} Resolution

In the search for Supersymmetry, the key observable (as described in Chapter 1) is the “missing transverse energy” (E_T^{miss}). A detector measurement of E_T^{miss} indicates the signal of the lightest supersymmetric particle, and in order to achieve better resolution for this observable, we will introduce new ECAL timing cuts.

Proton-proton (pp) collisions occur at the center of the CMS detector, when the Large Hadron Collider (LHC) collides two proton bunches running in the opposite directions. In 2011, the peak instantaneous luminosity reached $\sim 3.5 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$, and at that luminosity an average of 17 proton-proton collisions occurred in each bunch crossing. The extraneous proton-proton collisions that happen in the same bunch crossing as the collision of interest are referred to as “in-time pileup interactions”. The proton-proton bunch crossings occur at regular intervals and in 2011 the interval was 50 ns for most of the data-taking period. In the future this interval may decrease to 25 ns.

The detector (in particular the calorimeter) signals from pp collisions in bunch crossings earlier or later (“out-of-time pileup interactions”) often overlap and affect the detector signals for the pp collision of interest. Both in-time and out-of-time pileup pp interactions degrade the jet and missing E_T (E_T^{miss}) measurement resolutions (see e.g. Reference [16]). The degradation worsens when the proton beam intensity and instantaneous luminosity increase as well as when the bunch crossing interval shortens. In general, the resolution degrades with increasing number of vertices, and even with 1 vertex the resolution is degraded by ~ 6 GeV in quadrature with E_T^{miss} RMS values in 2011 data compared to the 2010 data due to out-of-time pileup. However, the degradation from the out-of-time pileup can be reduced by uti-

lizing the timing information of the detector signals, for example, by cutting on the electromagnetic calorimeter (ECAL) hit timing or using the reduced time integration for the hadron calorimeter (HCAL) reconstruction.

Some timing cuts are already applied to the ECAL reconstructed hits (rechits) by default in CMSSW. Rechits are produced from the uncalibrated rechits, which access the information from the digitized signals to reconstruct amplitude and time. At the rechit level, a summary of the information extracted at the uncalibrated rechit level is provided in the form of flags, which tell the reliability of the information contained in a rechit and in particular have the information about out-of-time signals (kOutOfTime). For calorimeter towers (CaloTowers) used in calorimeter jets and E_T^{miss} reconstruction, the ECAL rechits with kOutOfTime flags (set to false when $> 5\sigma$ away from 0 ns) are already removed in both ECAL Barrel (EB) and ECAL Endcap (EE) starting with the CMSSW_4.2.X release with an energy threshold of 2 GeV. For the particle flow reconstruction (see Section 4.1), the kOutOfTime cuts have been applied only in the EB with a threshold of 2 GeV. The existing time cuts are mainly for removing anomalous signals called “spikes” (see Section 4.1.2) purposes, but they also help out-of-time pileup mitigation [17].

Figure 3.1 shows the ECAL rechit time versus energy distributions, which are from the /L1Mu/Run2011B-14Oct2011-v1/RECO dataset collected in 2011 with 25 ns bunch spacings. The contributions from the out-of-time pileup interactions are evident in these distributions.

The main goal of the studies presented in this chapter is to determine how to apply timing cuts on the particle flow reconstruction in the ECAL end-cap, and how the cuts impact both jet and E_T^{miss} resolution. Additionally we determine whether or not the 2 GeV energy threshold cut in the ECAL Barrel is optimal, or if a more rigid cut would be beneficial.

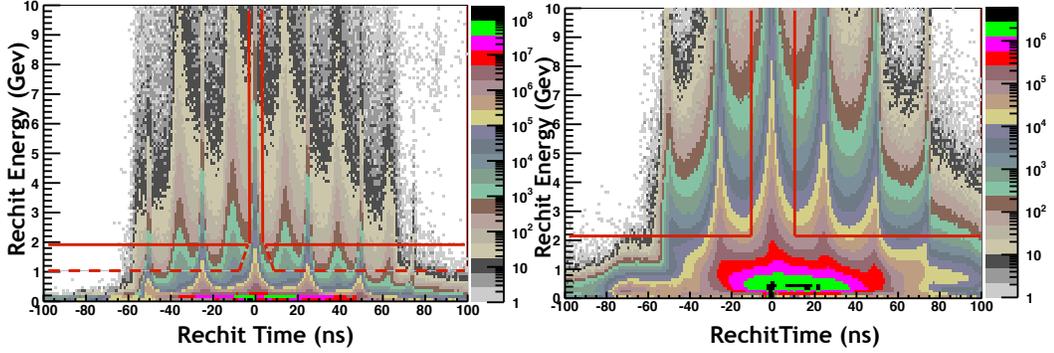


Figure 3.1: (Left) ECAL Barrel energy versus time distribution. The red lines show the 1 GeV and 2 GeV cut options. (Right) ECAL end-cap energy versus time plot. The only cut studied here is the 2 GeV threshold. In both distributions, the red lines are for illustration purposes and do not show accurate 1 and 2 GeV cut values for time, which vary as a function of energy.

In the following sections, we describe the analyses performed to provide the answers to these questions through rigorous jet and E_T^{miss} resolution studies using both Monte Carlo (MC) and 2011 collision datasets.

3.1 E_T^{miss} Resolution Studies

The ECAL rechit timing cuts are devised to improve the resolution of the E_T^{miss} measurement, and we need to validate the introduced changes extensively.

We began working with Monte Carlo (MC) samples to obtain baseline measurements of E_T^{miss} , its degradation with pileup interactions, and the effects of the timing cuts. Then we move onto a large primary dataset for jets to test the possible over-cleaning of the in-time signals in the actual collision data. Once this was completed we used samples collected with 25 ns bunch spacings in order to demonstrate the improvements in the E_T^{miss} reconstruction. The details of this analysis are described below.

3.1.1 Tests with the Neutrino Gun Pileup MC Samples

To better study the effects of pileup, a dataset constructed from a single neutrino (ν) is created and then combined with a set of proton-proton interactions. We do this so that there are no contributions from main collisions to complicate the analysis. Here, a simulation with 10 pp interactions were overlaid to single neutrino events to test in-time and out-of-time samples with both ± 25 ns and ± 50 ns delay. The MC samples used are on the Fermilab LPC cluster, under the `/uscms_data/d2/mikeh/work/out-of-time_samples/` directory:

- `SingleNu_RAW2DIGI_L1Reco_RECO_PU_10int_0ns.root` (In-Time Sample)
- `SingleNu_RAW2DIGI_L1Reco_RECO_PU_10int_m25ns.root` (25 ns Early Sample)
- `SingleNu_RAW2DIGI_L1Reco_RECO_PU_10int_p25ns.root` (25 ns Late Sample)
- `SingleNu_RAW2DIGI_L1Reco_RECO_PU_10int_m50ns.root` (50 ns Early Sample)
- `SingleNu_RAW2DIGI_L1Reco_RECO_PU_10int_p50ns.root` (50 ns Late Sample)

As shown in Figure 3.2, there are significant contributions from ± 25 ns bunch spacings, and much smaller contributions from the ± 50 ns bunch spacings. Figures 3.3 and 3.4 show E_T^{miss} , $\sum E_T$, $E_T^{\text{miss}}x$, and $E_T^{\text{miss}}y$. For $E_T^{\text{miss}}x, y$ the mean is ~ 0 , and the RMS is a good indicator of the intrinsic E_T^{miss} resolution. The statistics boxes for the $E_T^{\text{miss}}x, y$ distributions show that there is a small but clear decrease in root-mean-square (RMS) between the default configuration described in the introduction and the additional timing cut based on the `kOutOfTime` flag for EE rechits with energy > 2 GeV. Also shown is the application of the `kOutOfTime` removal in EB down to 1 GeV, resulting in additional reductions in the RMS value.

Tables 3.1 and 3.2 are summaries of Figures 3.3–3.4 along with results from the in-time and ± 50 ns out-of-time pileup MC samples. These show the $E_T^{\text{miss}}x(y)$ projection RMS values and difference in quadrature with and without the additional timing cut. Cut (1) refers to the timing cut with the energy threshold of 2 GeV in EB and 2 GeV in EE. Cut (2) refers to the timing cut with the energy threshold of 1 GeV

in EB and 2 GeV in EE. There is a small but clear decrease in $E_T^{\text{miss}}x(y)$ projection RMS values between the default configuration and the additional kOutOfTime-flag-based timing cut for EE rechits with energy > 2 GeV (cut 1) in the case of the out-of-time pileup samples. This indicates that the additional timing cut will reduce the E_T^{miss} resolution degradation from out-of-time pileups. Another test of applying the kOutOfTime removal in EB down to 1 GeV gives some additional smaller reduction in RMS values. In general, this leads us to expect a 1–2 GeV reduction in E_T^{miss} resolution at 10 out-of-time collisions. The left-hand panel of Figure 3.1 shows the contributions from primary and out-of-time collisions for a 1 or 2 GeV threshold cut.

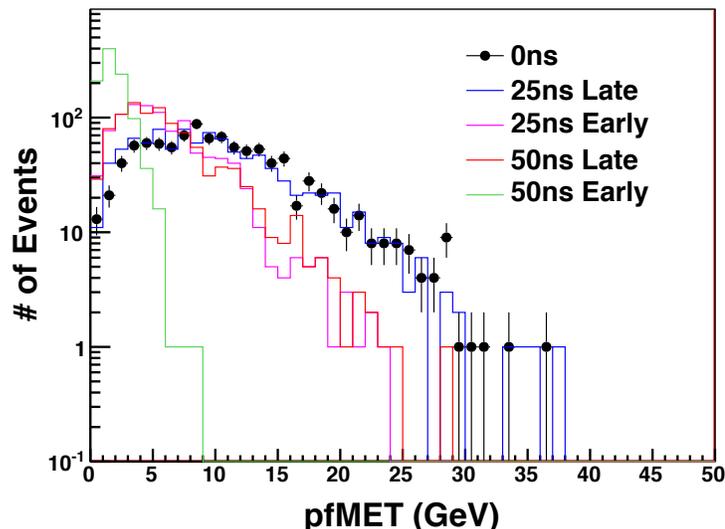


Figure 3.2: Particle flow E_T^{miss} distributions for the neutrino gun MC samples with 10 in-time or out-of-time (± 25 and ± 50 ns) pileup pp collisions.

3.1.2 Tests with the pp Collision Data

The MC studies presented above suggest that the additional timing cut will help remove the out-of-time pileup effects and improve the E_T^{miss} resolution. Now we test the timing cuts in the collision data sample using the Jet primary data set: /Jet/Run2011A-05Aug2011-v1/RECO.

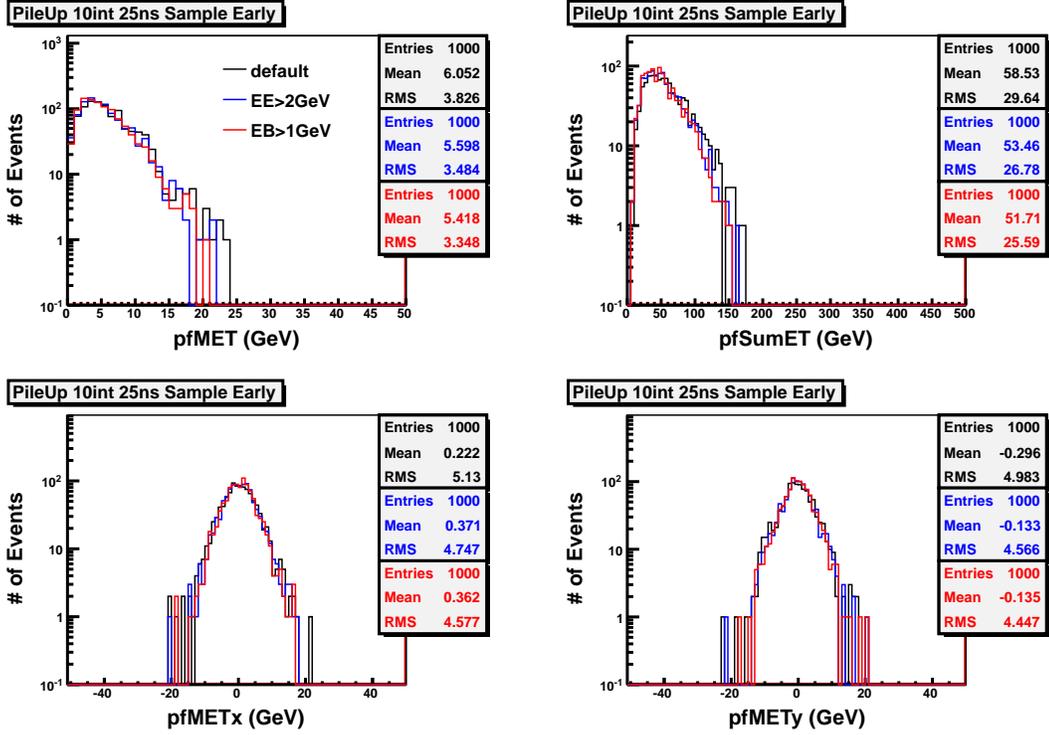


Figure 3.3: (25 ns Early Sample) Top row: (Left) Particle flow E_T^{miss} , (Right) Particle flow ΣE_T . Bottom row: (Left) Particle flow E_T^{miss} projected in the x direction, (Right) Particle flow E_T^{miss} projected in the y direction.

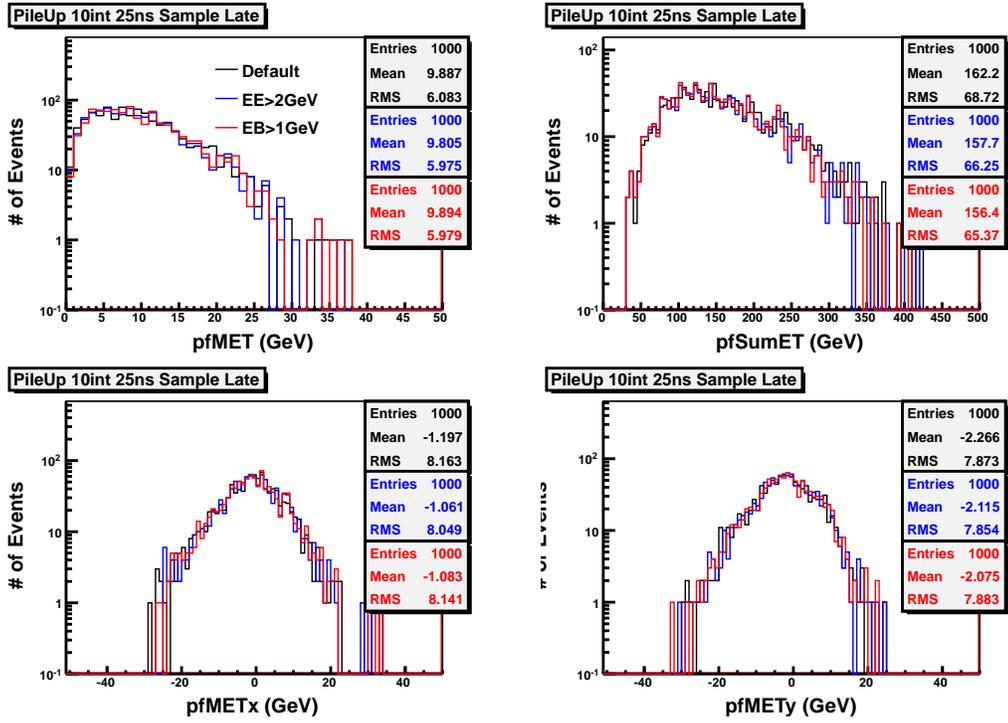


Figure 3.4: (25 ns Late Sample) Top row: (Left) Particle flow E_T^{miss} , (Right) Particle flow ΣE_T . Bottom row: (Left) Particle flow E_T^{miss} in the x direction, (Right) Particle flow E_T^{miss} in the y direction.

Table 3.1: Particle flow E_T^{miss} x projection RMS values in GeV with the default and two different additional timing cuts (1 and 2) discussed in the text, and the quadratic differences with and without the additional cuts.

Sample	Default	Cut (1)	Diff in Quad (1)	Cut (2)	Diff in Quad (2)
0 ns	8.74	8.74	0.00	8.74	0.00
25 ns Early	5.12	4.47	1.96	4.58	2.32
25 ns Late	8.16	8.05	1.45	8.14	0.69
50 ns Early	1.64	1.59	0.52	1.53	0.67
50 ns Late	5.45	5.36	0.99	5.36	1.04

Table 3.2: Particle flow E_T^{miss} y projection RMS values in GeV with the default and two different additional timing cuts (1 and 2) discussed in the text, and the quadratic differences with and without the additional cuts.

Sample	Default	Cut (1)	Diff in Quad (1)	Cut (2)	Diff in Quad (2)
0 ns	8.36	8.36	0.13	8.36	0.18
25 ns Early	4.98	4.57	2.00	4.45	2.25
25 ns Late	7.87	7.85	0.53	7.88	0.13
50 ns Early	1.60	1.47	0.61	1.43	0.70
50 ns Late	5.16	5.05	1.14	5.03	1.19

In the left panel of Figure 3.5, we compare the particle flow cluster (PFcluster) E_T^{miss} with and without the additional kOutOfTime timing cut for EE (cut 1) on an event-by-event basis. For a detailed description of (PF) E_T^{miss} see Chapter 4. Three different configurations were used during the analysis: kOutOfTime, default, and an additional ECAL Barrel cut. The default configuration refers to a timing cut in the ECAL barrel for energy > 2 GeV and no cut on the ECAL end-cap, the kOutOfTime refers to the timing cut in both ECAL barrel and end-cap regions with the energy threshold > 2 GeV, and the last configuration uses the timing cut with the 1 GeV and 2 GeV energy thresholds for EB and EE respectively. The plot shows (PFcluster) E_T^{miss} (kOutOfTime setting) subtracted from (PF) cluster E_T^{miss} (default setting). Most of the events in the Jet primary dataset do not have intrinsic E_T^{miss} ;

therefore, the E_T^{miss} decrease by the additional cut is a good indication of the removal of the anomalous signals. The E_T^{miss} increase is an indication of the over-cleaning of the real in-time hits, which we would like to avoid. The left panel of Figure 3.5 shows a comparison of the (PFcluster) E_T^{miss} with and without the kOutOfTime timing cut event-by-event. It is clear from this that we may be removing good reconstructed hits that have been flagged as out-of-time.

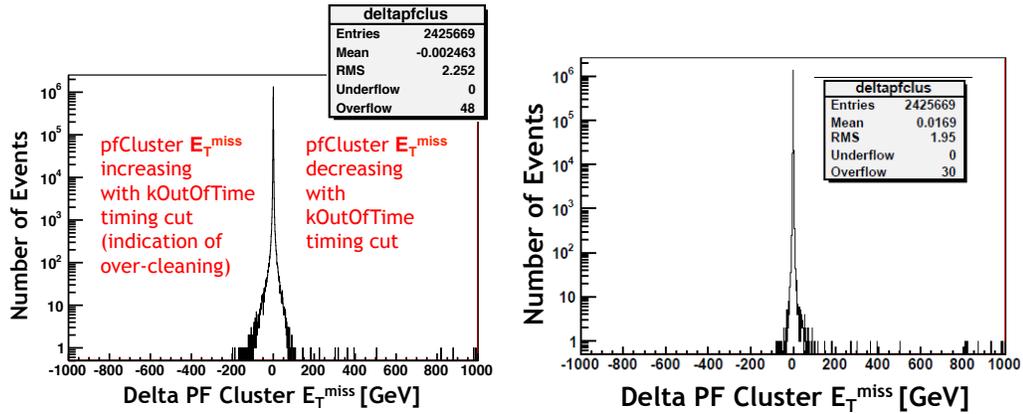


Figure 3.5: Particle flow cluster E_T^{miss} difference between default and kOutOfTime settings. (Left) Before time cut implementation showing large low E_T^{miss} tail due to over cleaning. (Right) After time cut applied showing reduced low E_T^{miss} tail.

We then scanned tail events that had high (16 events) and low (53 events) differences in (PFcluster) E_T^{miss} to determine if any of these events were actually “real” in-time events. The time distribution of kOutOfTime flagged reconstructed hits (Figure 3.6) with energy > 50 GeV¹ (most events looked like dijet events in a fireworks display Figure 3.7) and those that decreased by more than 600 GeV² were collected in order to identify any events which may be of concern (most do not look like real events in the fireworks display). See Chapter 4 for a detailed

¹ http://hep.baylor.edu/hatake/phys/tmp/pickevents_merged428H.root

² http://hep.baylor.edu/hatake/phys/tmp/pickevents_merged428L.root

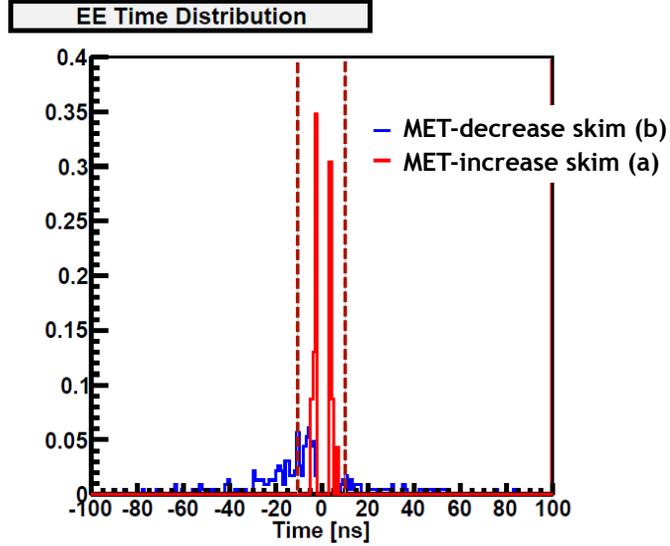


Figure 3.6: Normalized time distribution of EE rechits for a E_T^{miss} -decreased and increased skim sample (see text for details).

description of real and anomalous events in fireworks displays. In our sample, one event was removed that had hundreds of kOutOfTime flagged hits to get a better indication of how a standard sample would behave.

Due to the time distribution of the skim samples, we determined that it might be safer to keep all the reconstructed hits within 10 ns of the collision point whether or not they are flagged as kOutOfTime. Afterwards, we added the timing cut where $|t| > 10$ ns, and removed only the kOutOfTime reconstructed hits that exceeded this time cut. By comparing the left and right panels in Figure 3.5, it is evident that the over-cleaning of in-time rechits reduced significantly, as the low tail in the (PFcluster) E_T^{miss} difference distribution decreased.

Finally, to test the new configuration with the kOutOfTime flag and the $|t| > 10$ ns cuts, we obtained a Fall 2011 dataset that tested the 25 ns bunch spacings.

- /L1Mu/Run2011B-14Oct2011-v1/RECO

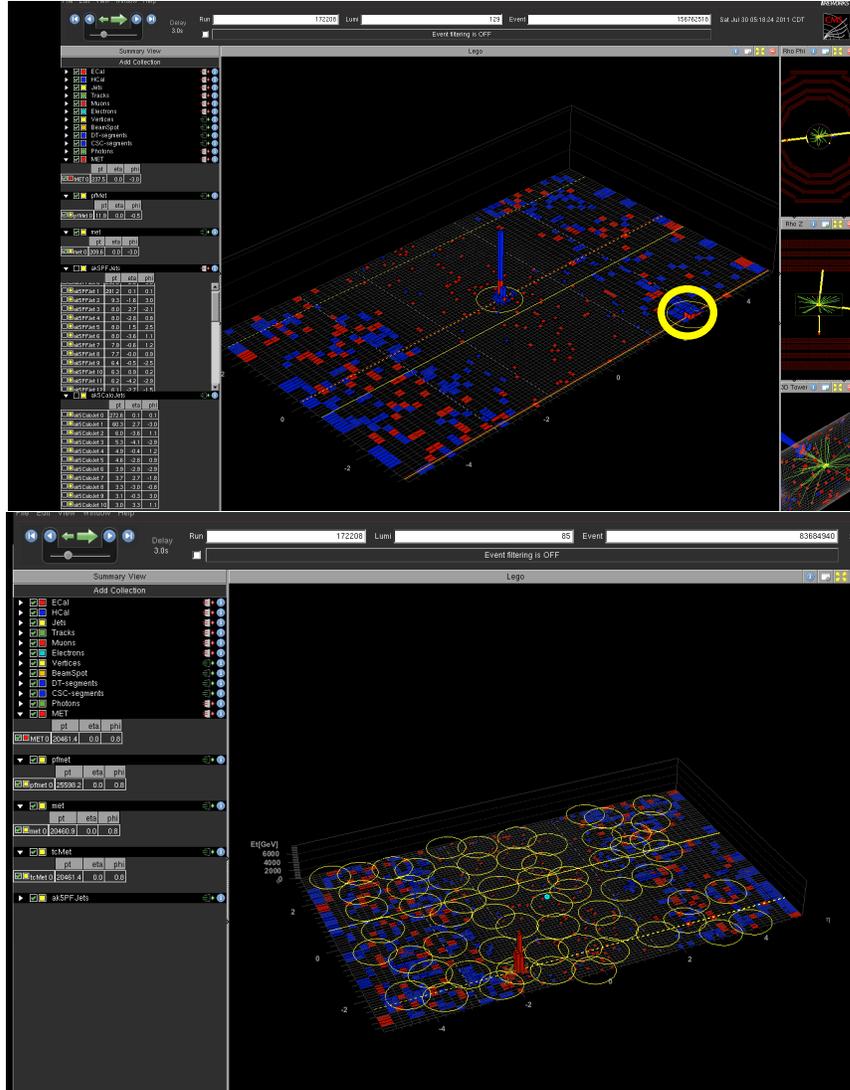


Figure 3.7: Fireworks display for both high and low Δ (PFcluster) E_T^{miss} . More events are described in Reference [18]. The locations of the skim samples are provided in the caption of Figure 3.6. (Top) Single sample event (172208:129:156762518) where particle flow cluster E_T^{miss} increases by more than 150 GeV. (Bottom) Single sample event (172208:85:83684940) where particle flow cluster E_T^{miss} decreases by more than 600 GeV.

- /L1EG/Run2011B-14Oct2011-v1/RECO

The results of this test are shown in Figures 3.8–3.9. Specifically, in the statistics box of Figure 3.8, the total (PF) E_T^{miss} (top left) shows an RMS of 10.22 GeV from the default configuration and then 10.05 GeV when the timing cut in EE is applied. Each of the distributions in these figures show E_T^{miss} resolution improving with the additional timing cut in EE.

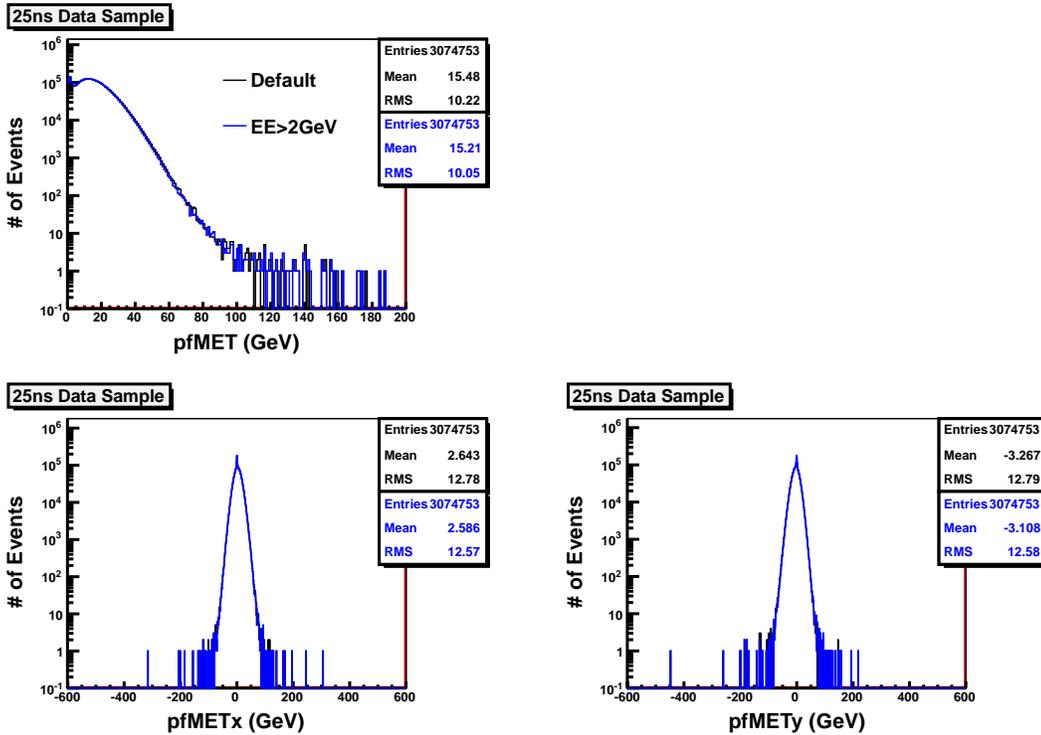


Figure 3.8: L1Mu 25 ns dataset tested with kOutOfTime flag and $|t| > 10$ ns. Top row: (Left) Particle flow E_T^{miss} . Bottom row: (Left) Particle flow E_T^{miss} in the x direction, (Right) Particle flow E_T^{miss} in the y direction. Small reduction seen in all RMS values in each figure’s adjacent statistic box.

3.2 Jet p_T Resolution Studies

The E_T^{miss} resolution study provided conclusive results for an improved E_T^{miss} resolution, hence now we explore the impact of the previously described cuts (kOutOfTime with and without the additional timing cut) on jet resolution. Additionally,

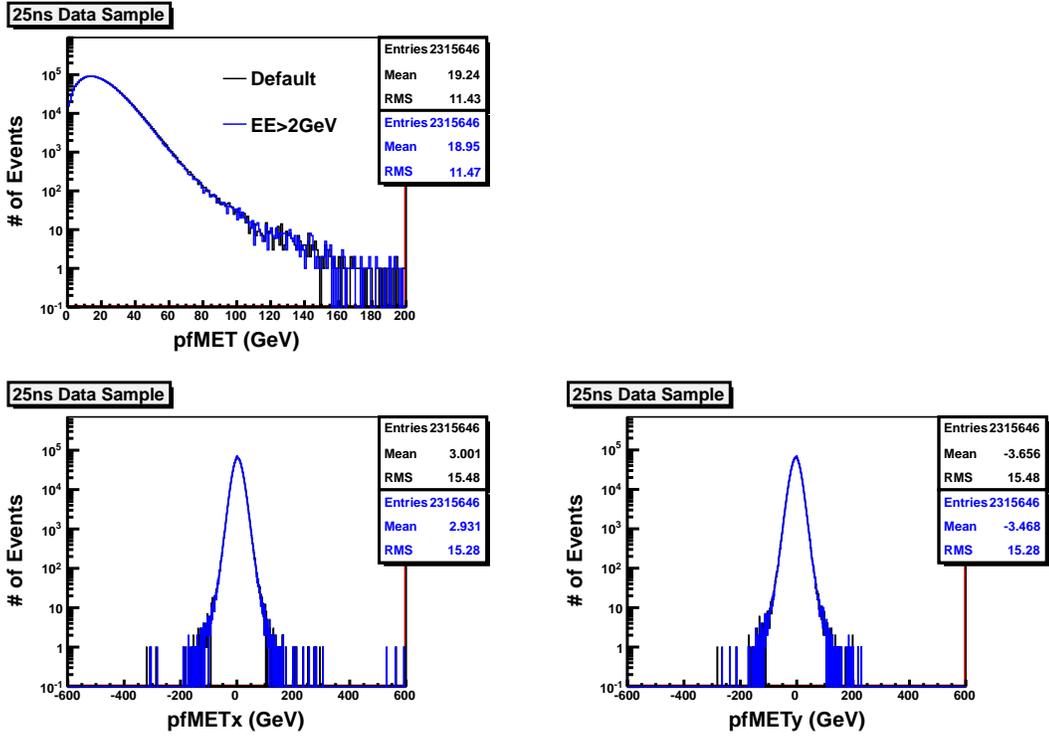


Figure 3.9: L1EG 25 ns dataset tested with kOutOfTime flag and $|t| > 10$ ns. Top row: (Left) Particle flow E_T^{miss} . Bottom row: (Left) Particle flow E_T^{miss} in the x direction, (Right) Particle flow E_T^{miss} in the y direction. Small reduction seen in all RMS values in each figure's adjacent statistic box.

the outcome is tested with a tighter time cut of $|t| > 7$ ns, where it is found there is no penalty to our current E_T^{miss} results. Three separate 25 ns bunch spacing datasets show improved RMS results after being tested. This tighter threshold of $|t| > 7$ ns was tested, but it was decided to commit the $|t| > 10$ ns into the final software modification to avoid any unanticipated effects of potentially removing in-time rechits. Figures 3.10–3.14 present the results from the following datasets:

`/DYToMuMu_M-20_TuneZ2_7TeV-pythia6/Summer11-HCa125_2TS_PU_S4_START42_V11-v2`

`/G_Pt-15to3000_TuneZ2_Flat_7TeV_pythia6/Summer11-HCa125_2TS_PU_S4_START42_V11-v2`

When we study jet p_T resolution in Drell-Yan events, we select only jets away from muons coming from Z decays. We use jets clustered by the anti- k_T algorithm with the size parameter $D = 0.5$, thus if it is found that jets have muons with $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \leq 0.5$ in η - ϕ space (see Figure 2.6), those jets are not used for jet p_T resolution studies.

We check that the generated jets have $|\eta| < 1.3$ to obtain results specifically for the barrel, or $1.5 \leq |\eta| \leq 3.0$ for results in the end-cap. If the distance between the generated jets and the particle flow jets is ≤ 0.1 , the jets are not used for this study.

Figures 3.10 and 3.12–3.14 show studies on the jet p_T resolutions in the end-cap region with and without the extra timing cut of removing ECAL rechits with energy above 2 GeV and $|t| > 7$ or 10 ns. The jet resolution is improved with the timing cut, (e.g. the (PF) Jet distribution in Figure 3.10 shows the RMS values for the default setting as 17.46 GeV and the additional timing cut as 17.31 GeV) and the 7 ns and 10 ns cuts do not make an appreciable difference, which leads us to use the conservative 10 ns cut.

Figure 3.11 indicates some improvement in the jet resolution by applying the timing cut to rechits with energy down to 1 GeV; however, the improvement is marginal. Thus, we suggest no change in this energy cut for the moment.

The ECAL timing cut presented here has been introduced into the 2012 data reconstruction software as the default configuration. This allows the entire CMS community to benefit from this improvement.

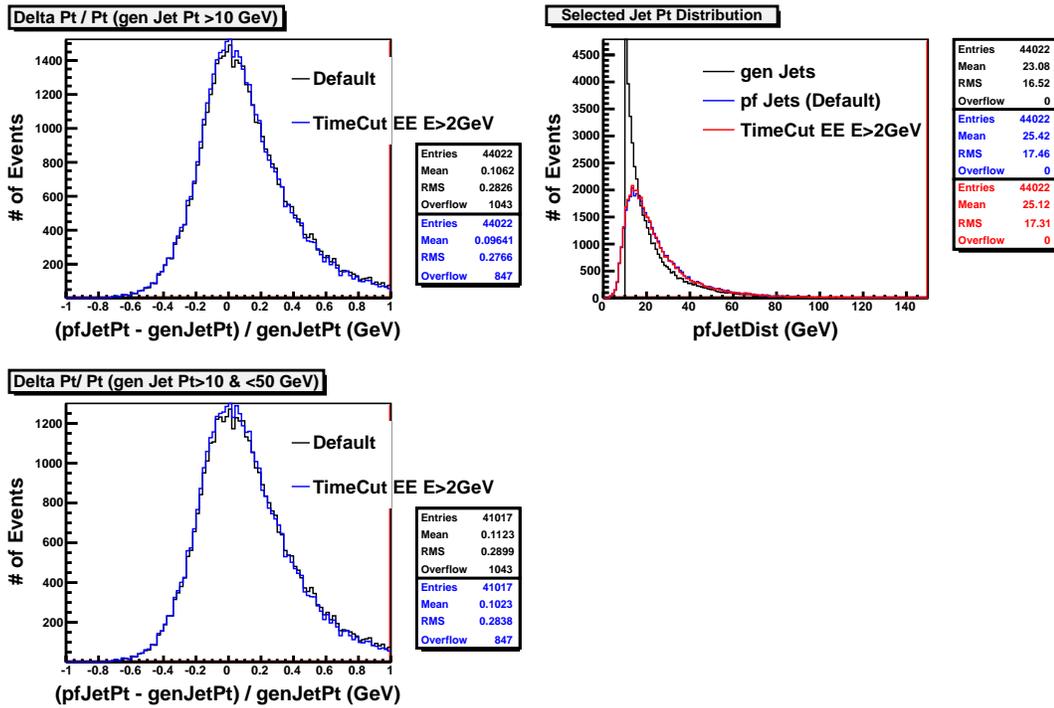


Figure 3.10: Jet p_T resolutions for jets in the ECAL end-cap region ($1.5 \leq |\eta| \leq 3.0$) with $p_T > 10$ GeV (Top Left) and with $10 < p_T < 50$ GeV (Bottom Left) with and without the timing cut, i.e. removing ECAL rechits with $|t| > 7$ ns and energy above 2 GeV in the Drell-Yan MC sample simulated with the 25 ns bunch spacing. The jet p_T distributions for the generator-level jets and particle flow jets are also shown (Top Right).

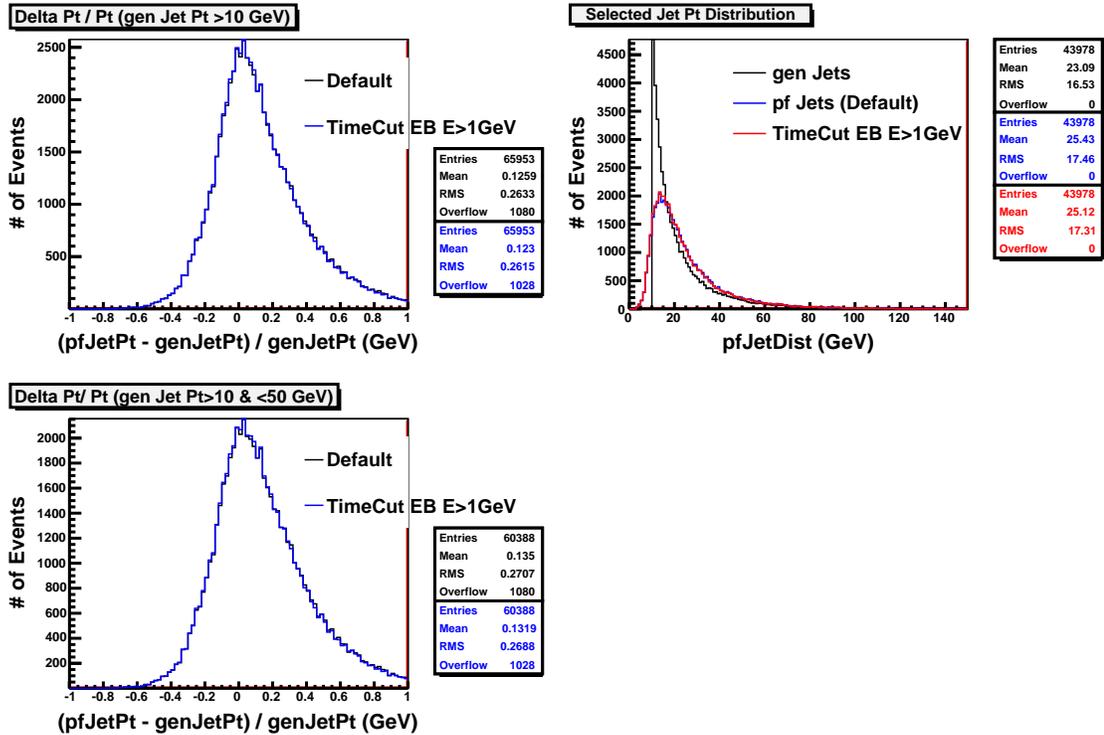


Figure 3.11: Jet p_T resolutions for jets in the ECAL barrel region ($|\eta| \leq 1.3$) with $p_T > 10$ GeV (Top Left) and with $10 < p_T < 50$ GeV (Bottom Left) with the timing cut applied for ECAL rechits down to 2 GeV (default) and to 1 GeV, in the Drell-Yan MC sample simulated with the 25 ns bunch spacing. The jet p_T distributions for the generator-level jets and particle flow jets are also shown (Top Right).

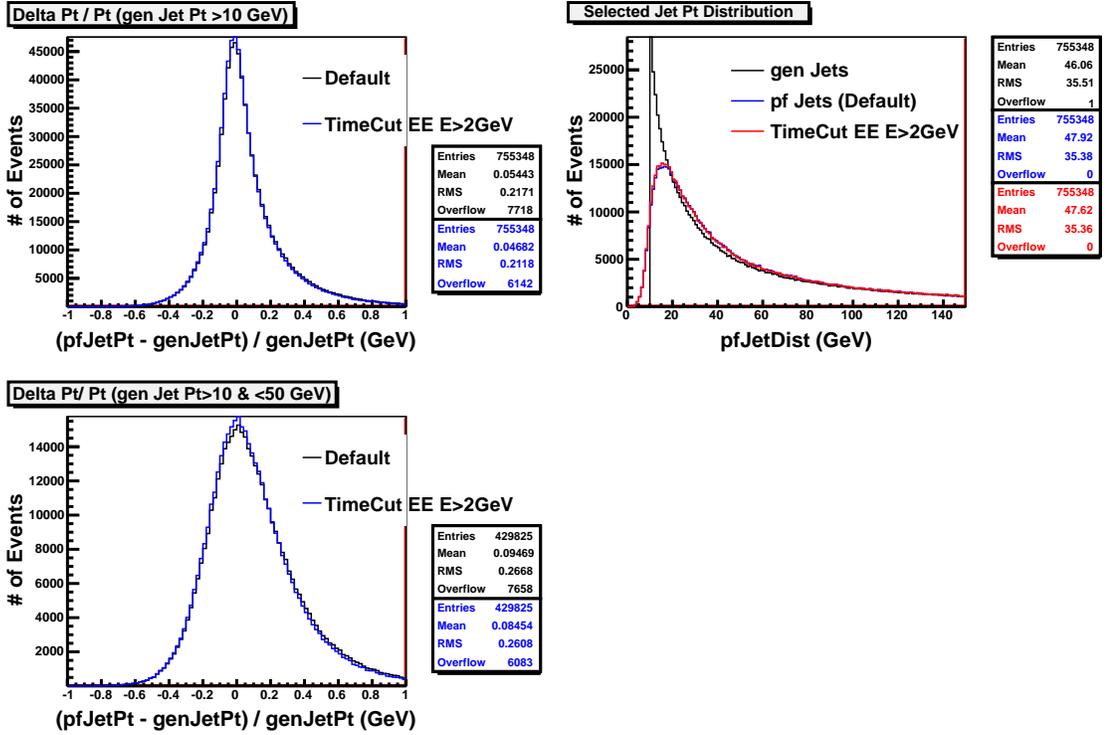


Figure 3.12: Jet p_T resolutions for jets in the ECAL end-cap region ($1.5 \leq |\eta| \leq 3.0$) with $p_T > 10$ GeV (Top Left) and with $10 < p_T < 50$ GeV (Bottom Left) with and without the timing cut, i.e. removing ECAL rechits with $|t| > 7$ ns and energy above 2 GeV in the photon plus jets MC sample simulated with the 25 ns bunch spacing. The jet p_T distributions for the generator-level jets and particle flow jets are also shown (Top Right). These are the same distributions as those in Figure 3.10 but the photon plus jets MC sample is used instead of the Drell-Yan MC sample.

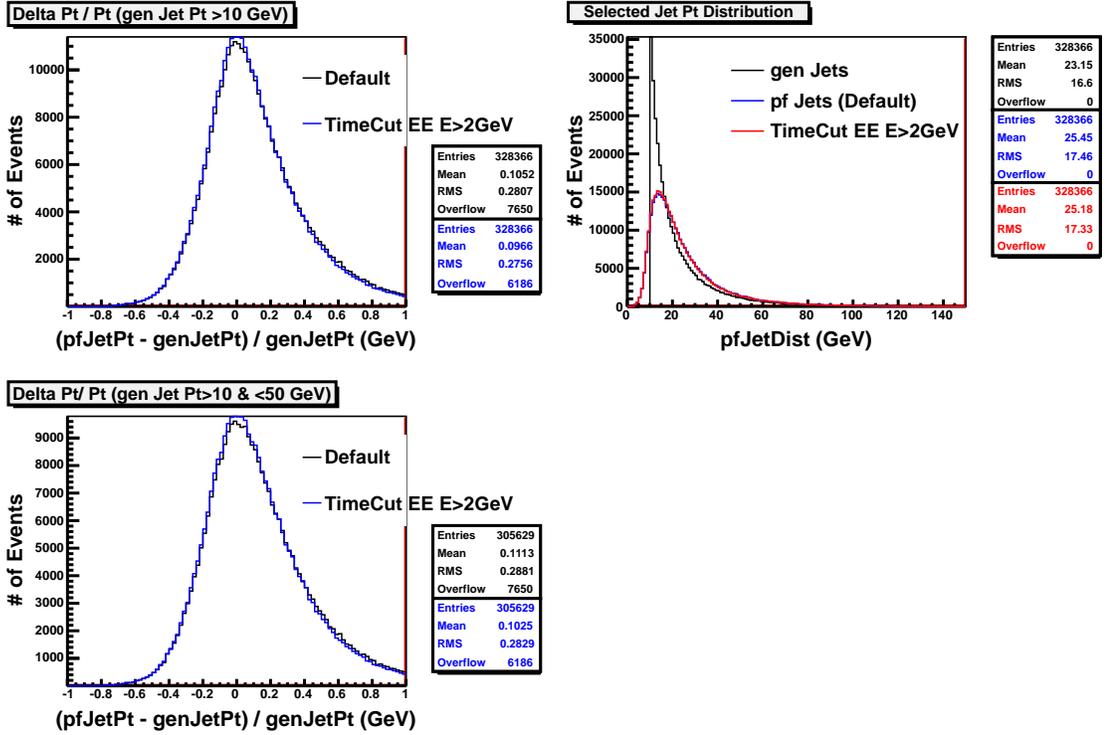


Figure 3.13: Jet p_T resolutions for jets in the ECAL end-cap region ($1.5 \leq |\eta| \leq 3.0$) with $p_T > 10$ GeV (Top Left) and with $10 < p_T < 50$ GeV (Bottom Left) with and without the timing cut i.e. removing ECAL rechits with $|t| > 10$ ns and energy above 2 GeV, in the Drell-Yan MC sample simulated with the 25 ns bunch spacing. The jet p_T distributions for the generator-level jets and particle flow jets are also shown (Top Right). These are the same distributions as those in Figure 3.10 but the timing cut is $|t| > 10$ ns instead of $|t| > 7$ ns to be safer against removing in-time rechits.

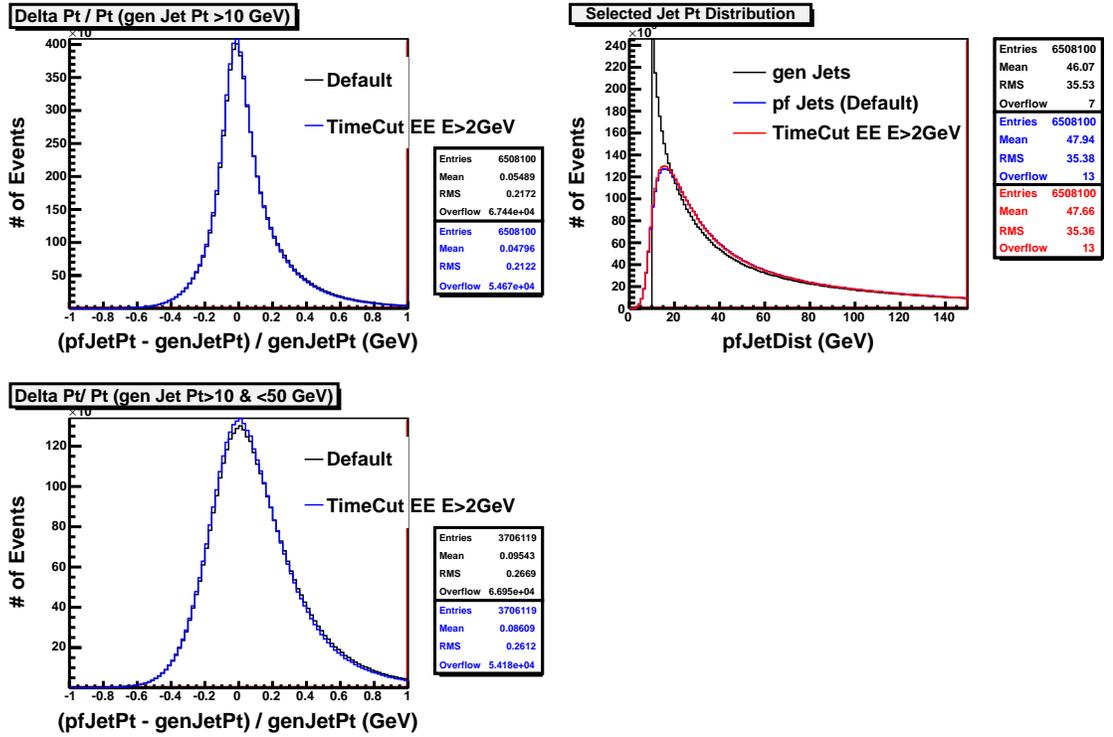


Figure 3.14: Jet p_T resolutions for jets in the ECAL end-cap region ($1.5 \leq |\eta| \leq 3.0$) with $p_T > 10$ GeV (Top Left) and with $10 < p_T < 50$ GeV (Bottom Left) with and without the timing cut i.e. removing ECAL rechits with $|t| > 10$ ns and energy above 2 GeV, in the photon plus jets MC sample simulated with the 25 ns bunch spacing. The jet p_T distributions for the generator-level jets and particle flow jets are also shown (Top Right). These are the same distributions as those in Figure 3.12 but the timing cut is $|t| > 10$ ns instead of $|t| > 7$ ns to be safer against removing in-time rechits.

CHAPTER FOUR

Search for Supersymmetry with Jets and E_T^{miss}

As discussed in Chapter 1, the R -parity conserving SUSY models predict the presence of E_T^{miss} events that help to identify the lightest supersymmetric particles (LSPs). There are, however, also standard model processes in which neutral weakly interacting particles such as neutrinos normally escape from the detector without producing any direct response in the detector elements, yielding high E_T^{miss} events.

The presence of such particles in collider experiments must be inferred from the imbalance of total momentum. The vector momentum imbalance in the plane perpendicular to the beam direction is particularly useful in hadron colliders, and is known as missing transverse momentum. Its magnitude is the missing transverse energy (E_T^{miss}). Up to this point we have considered events with genuine E_T^{miss} from either neutrinos or LSPs; now we will focus on the events that fall into the high E_T^{miss} samples but do not show true physics-like signals. The main sources of high E_T^{miss} events with “fake” E_T^{miss} fall into several categories: instrumental anomalous signals, beam-induced signals, and cosmic-ray backgrounds.

As the first step of this search, I conducted studies of high E_T^{miss} events which are presented in this chapter. Following that are details of the SUSY search including a description of how events are selected and a breakdown of four main sources of background contributions to our sample. Understanding the background contributions is crucial for determining if a SUSY signal is found in the data and will be used later for the simplified model analysis in Chapter 5.

4.1 E_T^{miss} Scanning

The main motivation for the 2012 CMS E_T^{miss} scanning group, led by Dr. Hongxuan Liu (Baylor University), is to perform a scan of 2012 collision data to identify sources of instrumental noise or potential problems in the event reconstruction that produce large artificial E_T^{miss} .

Historically, visual scans of high E_T^{miss} events has been proven to be a useful tool for finding problems/inconsistencies in the E_T^{miss} reconstruction as well as improving the noise-cleaning algorithms (see Section 4.1.1). Many interesting physics events show no obvious source of noise and have high E_T^{miss} . These are typically multi-jet events where the large fake E_T^{miss} is produced by jet energy mis-measurements or jets that fall in the boundaries between sub-detectors. The goal is to separate events that have signals coming from instrumental noise or beam-induced backgrounds from events with real E_T^{miss} . (We cannot avoid this type of mis-measurement, but we can estimate its contribution to a specific event sample, as described in the QCD background section later.) Visual scans of high E_T^{miss} events have identified several sources of anomalous noise in the calorimeters and machine-induced background. Collision events that contain one of these anomalies produce an imbalance in the reconstruction of missing transverse energy. This imbalance can produce large tails in the E_T^{miss} distribution.

E_T^{miss} is defined as the magnitude of the vector momentum imbalance in the plane perpendicular to the beam direction (see Equation (1.5)). CMS has developed three distinct algorithms to reconstruct E_T^{miss} : (a) calorimetry (Calo) E_T^{miss} based on calorimeter energies and calorimeter tower geometry; (b) track-corrected (TC) E_T^{miss} calculated by replacing the calorimeter tower energies matched to charged hadrons with their corresponding charged-track momenta; and (c) (PF) E_T^{miss} calculated using a complete particle-flow technique [19]. All physics objects used in the analysis are reconstructed in a consistent way using the particle flow algorithm, which identifies

and reconstructs particles produced from a collision. The particle flow algorithm will be the primary reconstruction method considered here because it provides the best E_T^{miss} scale and resolution.

4.1.1 Event Selection for E_T^{miss} Scanning

Skims of interesting high E_T^{miss} events from the HT, **Electron**, **Jet**, and **Muon** primary datasets (see Chapter 2: CMS Computing Model) are produced every two weeks. These datasets were distributed between the members of the E_T^{miss} scanning group, and my contribution involved only the HT dataset and thus will be the focus of the following discussion.

The following categories of interesting events (investigating occasional large difference between the E_T^{miss} from different reconstruction algorithms) are defined as [20, 21]:

- **Category I:** Large (PF) E_T^{miss}
- **Category II:** Large (PF) E_T^{miss} and (Calo) $E_T^{\text{miss}} / (\text{PF}) E_T^{\text{miss}} < 0.25$
- **Category III:** Large (Calo) E_T^{miss} and (PF) $E_T^{\text{miss}} / (\text{Calo}) E_T^{\text{miss}} < 0.25$

Category I typically contains events with HCAL noise, cosmic muons, and beam halo events. HCAL noise is an anomalous noise in the HB and HE subdetectors from the hybrid photodiodes (HPDs) — used to convert the scintillator light into electrical output — and the readout boxes (RBXs) that contain them. Beam halo noise stems from machine-induced secondary particles which are produced in showers initiated from collisions of the beam with residual gas inside the LHC vacuum chamber. Category II mainly contains events with mis-reconstructed particle flow muons and cosmic muons. Category III primarily consists of events with EB/EE spikes, which are identified as large reconstructed E_T^{miss} in isolated ECAL barrel (or end-cap) crystal. Another commonly found cause of high E_T^{miss} events is a tracking failure. The tracking failure filter rejects two sources of mis-reconstructed events

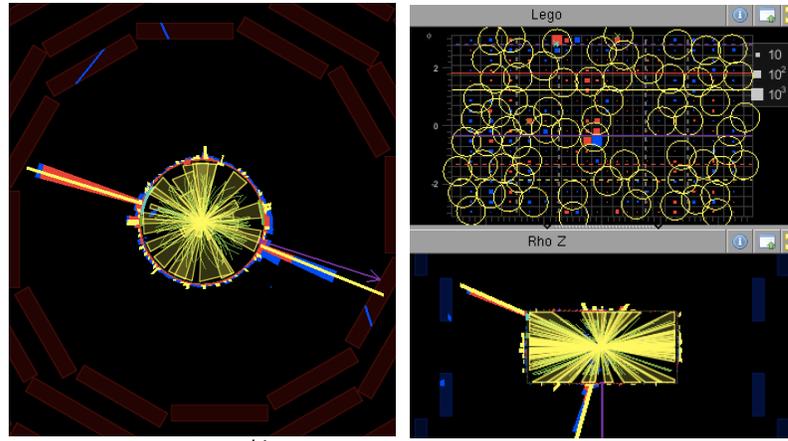
where standard or large calorimetry deposits contrast with a lack of reconstructed tracks.

These cuts evolved through the 2011 data-taking period to ensure that a reasonable number of events were used for thorough scanning. When anomalies are found, scanners communicate to detector reconstruction software experts to help determine necessary improvements to noise cleaning filters. When updated tools become available, scanners test them on skimmed high E_T^{miss} events to determine the efficiency of newly developed software (e.g. newly developed event filter or reconstruction code) in order to ensure that the events are either properly removed or have no longer high fake E_T^{miss} [22].

The CMS software *Fireworks* [14] has been used to create the event displays and study the events. Fireworks illustrations will be used here to demonstrate several types of identified anomalies found during the visual scan. For clarity, Figure 4.1 depicts a signal event from the scan.

191247:93:145914921

Category 2: Top 22 events with pfMET>300 && caloMET/pfMET<0.25



	pt,	eta,	phi
CaloMET:	28.2	0.00	1.26
TCMET:	13.4	0.00	1.82
PFMET:	382.2	0.00	-0.31

• 3 Good events with no flags (this event)

4

Figure 4.1: Event display of a real physics-like event with high E_T^{miss} .

A real physics-like E_T^{miss} event as shown in Figure 4.1 satisfies at least the following criteria:

- (1) All jets have normal particle content (see Section 1.3 for details).
- (2) None of the jets fall into the HCAL problematic region ($|\eta| = 1.3$, where the HCAL barrel ends and HCAL end-cap begins; see Figure 2.6).
- (3) None of the jets coincide directly with a masked¹ ECAL tower.
- (4) One good pile-up vertex can be identified.

4.1.2 Results from High E_T^{miss} Tail Scanning

Upon initial scan of the first golden JSON (Java Script Object Notation)² file created for 2012 scans (Cert_190456-191859_8TeV_PromptReco_Collisions12_JSON.txt), 60 of 62 events in Category I were found to have ECAL spikes at the ECAL barrel/end-cap boundaries (see e.g. Figure 4.2) affecting all (Calo) E_T^{miss} , (TC) E_T^{miss} , and the (PF) E_T^{miss} . The dataset and CMS software version for this analysis are:

- Dataset: /HT/Run2012A-PromptReco-v1/AOD
- CMSSW release: CMSSW_5_2_3_patch3

the other datasets (Electron, Jet, and Muon) are not discussed here, as mentioned previously.

It was found that there are still events that show an ECAL spike with no HCAL component. Events with these characteristics are not physics-like signals and indicate that our event filters are not working perfectly. These events are flagged by the `trackingFailure` filter and have very high E_T^{miss} , and some events have > 7 TeV

¹ The EB and EE have single noisy crystals which are masked during reconstruction. All masked cells make up for only 1% of the total crystals in the electromagnetic calorimeter.

² JSON files describe luminosity sections (a lumi section is 23 seconds of CMS data) for the runs that are considered for good physics analysis.

191271:124:107918186

Category 1: Top 61 events with pfMET>1500

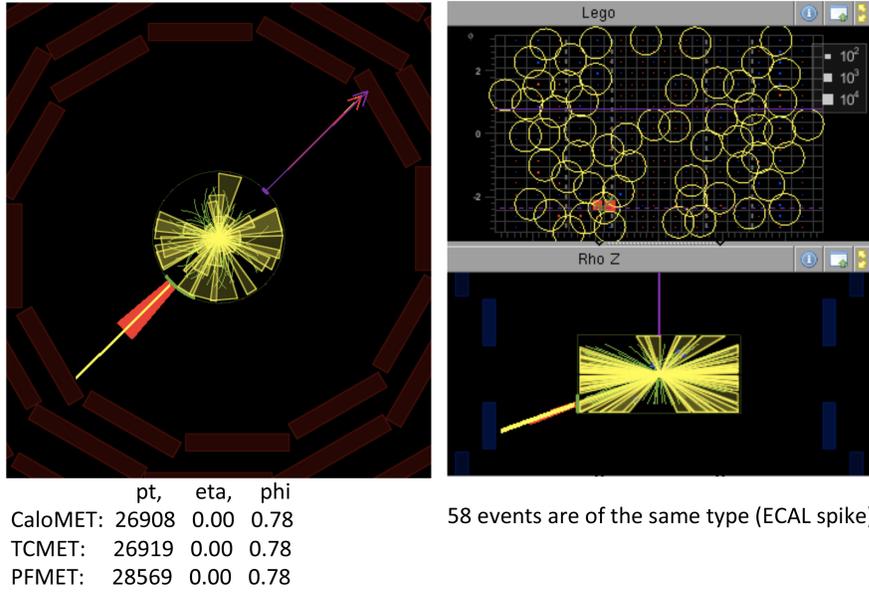


Figure 4.2: Event display showing an ECAL spike.

E_T^{miss} (e.g. 27 TeV). So for many of the events in this list, there is a problem in energy reconstruction. Since the energy deposition in the ECAL is so high in these events, the `trackingFailure` flag is being triggered all the time, but it is not believed to be related to any failure in the track reconstruction for these events.

It seems that there are two regions producing the spiky signals, and they seem to have problematic pulse shape, which may be related to an electronics failure. The high energy events appear to come from two supercrystals (5×5 crystal regions) in the EE. These supercrystals occasionally produce high amplitude anomalous pulses in several channels at once. Since they are not isolated, they are not caught by the `kWeird` cut active in the EE to remove the variable pulse throughout (VPT) discharge channels. In the ECAL cluster reconstruction, the information on the severity level of each potential seed rechit is calculated. Severity levels are available for objects to quantify the reliability of the energy information depending on the

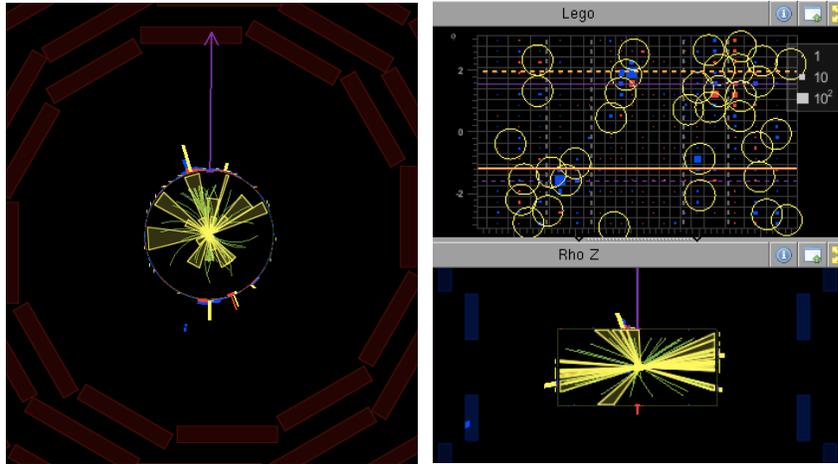
rechit. This in turn is compared to a list of severity levels which should be excluded. All rechits that register as aberrant channels (usually via the `kWeird` flag, which is the severityLevel corresponding to a spike) as defined in the configuration are removed.

Since they appear in single supercrystals, it seems plausible that they are related to some problem in the high-voltage power supply system, but this will need to be proven with further investigations on hardware, which we cannot access while taking data. Regarding a possible filter to remove events affected by this problem, we consider if the “discharges” are isolated to two particular supercrystals. This consideration implies that hits/jets are rejected if the two separate supercrystals affected are located in the regions: (ix: 21–25 iy: 21–25 zside: -1) (ix: 46–50 iy: 96–100 zside: +1). Most (>90%) of the events are found in the first supercrystal. A filter was developed by the ECAL software experts to remove these events.

There are two other types of anomalous events found in Category II and Category III. The top image of Figure 4.3 exhibits a high (PF) E_T^{miss} with lower (TC) E_T^{miss} and (Calo) E_T^{miss} due to an anomalously high PF muon. The ECAL dead cell filter (`ECALDeadCellBE`) refers to an event in which jet or particle information registers as a single noisy (or dead) EB or EE crystal that has been masked during reconstruction. This flag is not the source of the large (PF) E_T^{miss} , but is simply another indication that the high E_T^{miss} measurements may not be real. The bottom image of Figure 4.3 shows a E_T^{miss} event that is tagged by the `inconsistentMuon` flag. The inconsistent muon filter rejects events containing at least one particle-flow muon with $p_T > 100$ GeV identified as being both a tracker and a global muon, and with a tracker and combined fit momentum differing by more than 10%. These two types of anomalies account for a much smaller fraction of events and were noted but not addressed during this work.

191226:747:1018943227

Category 2: Top 22 events with $pfMET > 300$ && $caloMET/pfMET < 0.25$

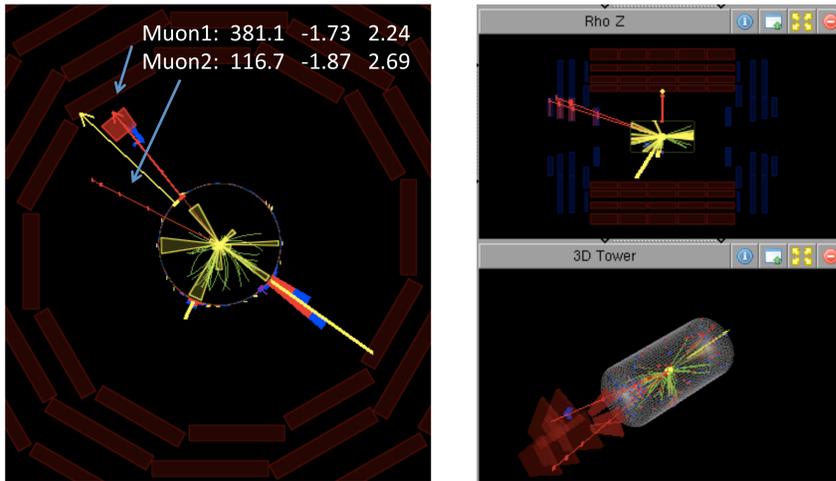


	pt,	eta,	phi
CaloMET:	91.3	0.00	-1.18
TCMET:	108.2	0.00	-1.22
PFMET:	957.0	0.00	1.58

Single event with flag:
['ECALDeadCellBE']
Due to high pt (~1200GeV) pfMuon at eta=-2.30

191226:1315:1632889079

Category 3: Top 72 events with $caloMET > 300$ && $pfMET/calMET < 0.25$



Muon1: 381.1 -1.73 2.24
Muon2: 116.7 -1.87 2.69

	pt,	eta,	phi
CaloMET:	453.3	0.00	2.39
TCMET:	365.6	0.00	2.25
PFMET:	26.1	0.00	-0.82

14 events with flag ['inconsistentMuon']

Figure 4.3: (Top) EcalDeadCell flagged event and (Bottom) an inconsistent-muon flagged event.

In conclusion, it was decided that the HT dataset would be divided into 3 separate subsamples in order to better scrutinize this particular sample. Updated ECAL filters were put into place along with new tests and macros developed for diagnostic plots for the dataset `/HTMHT/Run2012B-PromptReco-v1/RECO`.

The same type of analysis was performed on the new dataset and events were scanned. Similar types of events, as presented earlier, were found and thus are omitted here.

4.2 Event Selection and Analysis Details

The quest to find answers to the most basic questions of our universe has pushed the scientific community to explore higher energy thresholds in particle physics where new physics and possibly Supersymmetry (SUSY) exist. Theorists have predicted many SUSY models in which new exotic particles are expected and would shed light on some of these questions. The specific criteria for the Supersymmetry search listed here and in the following section are developed by the “RA2” analysis group in CMS. Here, I provide an overview of the search. My specific contributions and the interpretation of the results are then discussed in detail in Chapter 5.

The data samples used in the analysis are collected using a trigger based on the scalar sum ($H_T = \sum_{i=1}^{N_{\text{jet}}} p_{T,i}$) of the transverse momentum (p_T) of all the jets reconstructed at the trigger level. Alternatively, we collect samples using a trigger based on both H_T and H_T^{miss} , where H_T^{miss} defines the magnitude of the negative scalar sum p_T of jets having $p_T > 30$ GeV and $|\eta| < 5$. The triggers are fully efficient for the baseline event selection, where the baseline is defined as events with $H_T > 500$ GeV and $H_T^{\text{miss}} > 200$ GeV. All events have at least one good reconstructed pp interaction vertex, and have at least three jets with $p_T > 50$ GeV and $|\eta| > 2.5$.

Events with isolated muons and electrons with $p_T > 10$ GeV are removed in order to reject background events with leptons in the final state.

Jets are clustered using the anti- k_T algorithm (see Section 1.3) with size parameter $D = 0.5$. The raw jet momenta are corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in p_T . Contributions from additional pp collisions overlapping with the event of interest (pileup collisions) are mitigated using the Fastjet³ tools.

In order to be sensitive to different kinematics of the signal events in the sample, it is divided into 14 sub-samples defined in terms of the event H_T and H_T^{miss} values.

4.3 Standard Model Background Contributions

Standard model background events passing these requirements originate from several different sources, with primary contributions from [23, 24]: (1) $Z \rightarrow \nu\bar{\nu}$, (2) $t\bar{t}/W \rightarrow e, \nu + X$, (3) $t\bar{t}/W \rightarrow \tau_{\text{hadr}} + X$, and (4) QCD multijets. Other standard model processes that contribute to the background are found to be negligible.

To estimate these backgrounds, several Monte Carlo (MC) simulated samples are used to model the signal and validate the data-driven background prediction methods. The MC samples used are generated using MADGRAPH⁴ [25] and PYTHIA6 [26]. The samples are then processed with detector simulation software (Geant4-based⁵ CMS simulator), and the sample contains the same distribution of pileup pp collisions as in the data. After the samples are generated and run through

³ Software providing efficient geometrically-based implementations for the anti- k_T jet algorithm. It also provides access to tools that allow one to determine the areas of individual jets, which is of importance when correcting for the underlying event and pileup contamination.

⁴ Monte Carlo simulation program that allows you to generate amplitudes and events for any process (with up to 9 external particles) in any model.

⁵ Detector-specific software that simulates the particle interaction in the detector.

the baseline selection, 1,927 events remain, which is in reasonable agreement with what is observed directly in the data (see Table 4.1). Precise background contributions are an essential component when determining the 95% confidence level on the upper limit of the cross section for SUSY particles (see Chapter 5). These values, however, are not used for further calculations due to difficulties in evaluating event generation and simulation uncertainties. It should be noted that the jets + $E_{\text{T}}^{\text{miss}}$ analysis tries to estimate the yields from all possible standard model sources to each of the signal regions as discussed in the Chapter 4.12. After subtracting these from the data, an excess would indicate SUSY. By underestimating the backgrounds, a “fake” SUSY discovery may emerge, whereas overestimating will reduce the chances of an early discovery.

4.3.1 $Z(\nu\nu) + \text{Jets Background}$

The production of a Z boson and jets, followed by the decay of the Z boson into neutrinos, constitutes an irreducible background. Irreducible physics background processes involve two neutrinos which are invisible in the detector components. For other background processes, we can use a combination of kinematic properties and particle decay products in order to reduce their contributions. In the next section, we discuss how we suppress the background contributions from $W(l\nu) + \text{jets}$ and $t\bar{t}$ production, and also how we estimate the remaining backgrounds. Here, however, the neutrinos from the Z boson decay are invisible in the detector; therefore, we cannot reduce the background based on the Z decay products. The kinematic properties and hadronic components are similar for both the Z boson and photon at high p_{T} . The cross section ratio between Z -boson and photon production provides a robust prediction of the missing transverse momentum spectrum for invisible Z bosons at high p_{T} .

4.3.2 W and Top Background

In this search, events with isolated muons and electrons are vetoed by the selection requirements in order to suppress backgrounds from $W + \text{jets}$ and $t\bar{t}$ production in which W decays to a lepton and a neutrino. However, some of $W + \text{jets}$ and $t\bar{t}$ events are not vetoed when an electron or muon from a W decay is not identified, non-isolated, or outside the geometric and/or kinematic acceptances, or when the lepton from the W decay is a tau that decays hadronically. The first three cases are denoted as a “lost lepton” ($t\bar{t}/W \rightarrow e, \nu$) where the lepton can be out of acceptance because either its transverse momentum is too small, or it is emitted in the forward direction. This background is estimated from a $\mu + \text{jets}$ control sample that uses the same criteria defined for our search but requires exactly one isolated μ (rather than zero isolated μ). The events are then weighted according to the reconstruction efficiency to account for non-isolated or unidentified leptons. The hadronically decaying tau leptons ($t\bar{t}/W \rightarrow \tau_{\text{hadr}}$) are an important second contribution of the W and $t\bar{t}$ backgrounds. The estimation of this background contribution is similar to the leptonic decay, where a $\mu + \text{jets}$ control sample is selected from data. The hadronic properties of events in this background are identical to those of a muon control sample, except for the fraction of the τ jet energy deposited in the calorimeters, so the muons in the sample are replaced with a hadronically decaying τ . Selection cuts described before are then applied, as well as a correction for the relative branching fractions of W decays into muons or hadronic τ jets.

4.3.3 QCD Background

Finally, the QCD multijet background is estimated using a “rebalance-and-smear” method [24] directly from the data. This method predicts the full kinematics in multijet events, while being unaffected by events with true $E_{\text{T}}^{\text{miss}}$. Conceptually, this method is based around the idea that large missing transverse momentum arises

in QCD multijet events when one or more jets in the event have a jet energy response far from unity. The jet energy response is defined as the ratio of the transverse momentum of the reconstructed jet over the one which would result from measuring perfectly the momentum of particles in the jet. For jets with $p_T > 15$ GeV in the events, the p_T is adjusted using a kinematic fit such that the events are balanced in the transverse plane. The rebalanced events are then smeared using the jet resolution distribution and analyzed to determine the QCD multijet background.

4.3.4 Background Estimation Summary

All the background contributions including their uncertainties are summarized in Table 4.1. The standard model background predictions listed here are consistent with data, indicating that there is no significant amount of SUSY signal in the data used for this search. The measurement from data (the number of events passing 14 search selections), background predictions, and the efficiency of the signal passing the selection, are used as inputs for all cross section upper limit calculations that will be described in the following chapter.

Table 4.1: Event yields for different backgrounds for the 14 search selections together with the total backgrounds, as determined from the collision data, and the number of events observed in data. [23]

Selection (GeV)		Z	$t\bar{t}/W$	$t\bar{t}/W$	QCD	Total	Data
H_T	H_T^{miss}	$\rightarrow \nu\bar{\nu}$	$\rightarrow e, \mu + X$	$\rightarrow \tau_{\text{hadr}} + X$	multijet	background	
500–800	200–350	359 ± 81	327 ± 47	349 ± 40	119 ± 77	1154 ± 128	1269
500–800	350–500	112 ± 26	48 ± 9	62.5 ± 8.7	2.2 ± 2.2	225 ± 29	236
500–800	500–600	17.6 ± 4.9	5.0 ± 2.2	8.7 ± 2.5	0.0 ± 0.1	31.3 ± 5.9	22
500–800	>600	5.5 ± 2.6	0.8 ± 0.8	2.0 ± 1.8	0.0 ± 0.0	8.3 ± 3.2	6
800–1000	200–350	48 ± 19	58 ± 15	56.3 ± 8.3	35 ± 24	197 ± 35	177
800–1000	350–500	16.0 ± 6.7	5.4 ± 2.3	7.2 ± 2.0	$1.2^{+1.3}_{-1.2}$	29.8 ± 7.5	24
800–1000	500–600	7.1 ± 3.7	2.4 ± 1.5	1.3 ± 0.6	$0.0^{+0.2}_{0.0}$	10.8 ± 4.0	6
800–1000	>600	3.3 ± 1.7	0.7 ± 0.7	1.0 ± 0.3	$0.0^{+0.1}_{0.0}$	5.0 ± 1.9	5
1000–1200	200–350	10.9 ± 5.1	13.7 ± 3.8	21.9 ± 4.6	19.7 ± 13.3	66 ± 15	71
1000–1200	350–500	5.5 ± 3.0	5.0 ± 4.4	2.9 ± 1.3	$0.4^{+0.7}_{-0.4}$	13.8 ± 5.5	12
1000–1200	>500	2.2 ± 1.7	1.6 ± 1.2	2.3 ± 1.0	$0.0^{+0.2}_{0.0}$	6.1 ± 2.3	4
1200–1400	200–350	3.1 ± 1.8	4.2 ± 2.1	6.2 ± 1.8	11.7 ± 8.3	25.2 ± 8.9	29
1200–1400	>350	2.3 ± 1.5	2.3 ± 1.4	$0.6^{+0.8}_{-0.6}$	$0.2^{+0.6}_{-0.2}$	5.4 ± 2.3	8
>1400	>200	3.2 ± 1.8	2.7 ± 1.6	1.1 ± 0.5	12.0 ± 9.1	19.0 ± 9.4	16

CHAPTER FIVE

Interpretation of the Supersymmetry Search with “Simplified Models”

In Chapter 4 the search for SUSY in the jets and missing E_T final state was explored, but no evidence of SUSY was found. This is far from the end of the story for the Supersymmetry search, as theorists respond to a null result by simply modifying their models to still make them consistent with data.

When we have null results in a search for a new physics model, one of the most effective ways to deliver the results to the particle physics theorists is to present the “limits” on the physics model considered and its model parameters. This allows physicists to focus on the models and their parameter space which are not excluded. In this chapter, the limits and interpretation derived from this SUSY search using a simplified model are discussed.

The simplified model spectra (SMS) were originally designed for the purpose of creating a language in which experimentalists and theorists could describe new physics model signatures without the need for experimental details [27]. Generally speaking, the simplified models are developed as an effective Lagrangian describing a small number of new particles and their interactions. A topology refers to one specific set of production and decay mode, usually ending with a standard model particle, and each can be mapped onto at least one physics object (i.e. hadronic or leptonic signatures).

Depending on the desired signature, a number of SMS were developed to describe hypothetical particles and decay chains. Each has a different topological signature that describes production and decay in terms of particle branching ratios and probable mass range for the LSP. The choice of SMS is dominated by [28]:

- **Color** Which colored particles dominate production?

- **Mass** What is the mass hierarchy of the particle spectrum?
- **Degeneracy** Is there mass degeneracy?
- **W or Z** Are weak bosons involved in the decay chain?
- **Generation** Are the third-generation particles (t , b , or τ) considered special?

Each model helps place limits on different theoretical SUSY models through calculation of 95% confidence level exclusion limits on signal cross sections as a function of the mass parameters. One common requirement of all topologies is the presence of a “physical” source of missing transverse energy, which will be described in detail in the following sections.

5.1 Calculating Cross Section Upper Limits

When we do not find an evidence of a signal (i.e. there is no significant excess in data compared to the expected background from the known sources, as in the case discussed in Chapter 4), we often proceed to set a limit on the signal cross section. This will provide a information on “*if there is actually a signal hidden in our data, but too small an amount so that it is not visible in our current data, how large can this hidden signal be?*” This is generally called the upper limit on the signal cross section. By comparing these cross section upper limits with the signal cross section predicted by a theoretical model such as the simplified model, we may exclude a model completely or a certain parameter space of a model.

The basic idea on how to determine the signal cross section upper limit is based on simple Poisson statistics. Consider a simple case in which you are looking for a signal in a single channel and the average expected number of events in data is μ . The probability that we observe n events, given μ , is determined by a Poisson distribution as:

$$P(n|\mu) = \frac{\mu^n e^{-\mu}}{n!}$$

In case of multiple bins (channels) as described in Chapter 4, this can be extended to a joint probability:

$$P(\bar{n}|\bar{\mu}) = \prod_i^{\text{nbins}} \frac{\mu_i^{n_i} e^{-\mu_i}}{n_i!}$$

where \bar{n} denotes a set of observations n_i and $\bar{\mu}$ refers to a set of parameters μ_i .

The total expected number of events can have only a background contribution, $\mu_i = b_i$, for the probability of the observation in the background-only hypothesis (in the hypothesis that the data come only from the background). We can also define μ_i as $\mu_i = s_i + b_i$ in the signal + background hypothesis (in the hypothesis that the data include both the background and signal events). For simplicity, we denote the probabilities in the signal + background hypothesis and background hypothesis by $P(\text{data}|\text{signal} + \text{background})$ and $P(\text{data}|\text{background})$ hereafter.

There are different variants of the method to compute the limits, and in this analysis the CL_s method [29, 30] is adopted as recommended by the ATLAS and CMS statistics committees. In the CL_s method, a *test statistic* (Q) is defined by the likelihood ratio of the Poisson probabilities as:

$$Q = \frac{P(\text{data}|\text{signal} + \text{background})}{P(\text{data}|\text{background})} \quad (5.1)$$

to characterize the difference of the observation in the signal + background hypothesis and in the background only hypothesis.

As shown previously, the Poisson distribution is used here as a joint probability to run over all channels. The signal estimation s_i we described is dependent on a variety of variables such as the expected signal cross section, the decay branching ratios, the integrated luminosity, as well as the detection efficiency of the signal. The background estimation b_i may be also dependent on similar variables such as the standard model background cross section and selection efficiencies. In this search, the background b_i come directly from data as discussed in Chapter 4.

For efficiency, the test statistic can be expressed in a logarithmic form:

$$-\ln Q = \sum_{i=1}^{\text{nbins}} s_i - \sum_{i=1}^{\text{nbins}} n_i \ln \left(1 + \frac{s_i}{b_i} \right)$$

which is expressed by a sum of event weights, $w_i = \ln \left(1 + \frac{s_i}{b_i} \right)$. The signal to background ratio must be generated by significant Monte Carlo statistics or be processed through a rebinning or smoothing procedure in order for it to be finite for this calculation. In this way, a weight is assigned to each event and is dependent on the test variable.

The confidence level $\text{CL}_{\text{s+b}}$ defines the fraction of experiments in a large ensemble of experiments with the signal + background hypothesis which would produce results less signal-like than the observed data and can be written as:

$$\text{CL}_{\text{s+b}} = P(Q \leq Q_{\text{obs}} | \text{signal} + \text{background}) \quad (5.2)$$

where Q_{obs} is the value of the test statistic that was actually observed. This can be used to test the consistency of the data with the signal + background hypothesis. The probability in the upper tail of the Q distribution in the signal + background hypothesis may be used to exclude a signal hypothesis because it does not predict enough signal to explain the candidates in the data. Thus, if $\text{CL}_{\text{s+b}} < 0.05$, the signal + background hypothesis is excluded at the 95% confidence level.

The above discussions ignore the complications due to the systematic uncertainties in the estimated background and the amount of the signal passing the event selection. The treatment of systematic uncertainties in the limit calculation is rather complicated, and beyond the scope of this thesis. The uncertainties considered in the limit calculations are discussed in the subsequent sections. The detailed implementation of the CL_s method used in this analysis is reported in [31].

In summary, the signal hypothesis test using the CL_s method can be used as a function of the signal cross section in order to set a limit on the signal cross section at

a certain confidence level. The typical confidence level quoted is 95%. The 95% CL limit presented in the proceeding section is calculated using the method described here, and the algorithms that were already available in the CMSSW_4.2.8 release (developed for analysis of 2011 data) are used. Documentation¹ for the location and structure of the code for computing confidence limits using the method described here, along with various other methods, are available through the CMS TWiki pages for RooStats-based² statistic tools.

5.2 Topology Descriptions

As discussed above, the Supersymmetry search in the jets and E_T^{miss} channel discussed in Chapter 4 is interpreted here using the simplified model. Several topologies of the Supersymmetry signal are considered, which are discussed below.

The naming conventions for SMS follow a particular set of rules. The names always begin with the letter “T” (for topology), followed by a number and then a truncated description of particles involved in the decay chain. If the number is even, the topology represents squark-squark production, and if odd, gluino-gluino production. Here, the squark-gluino pair production topology has been omitted for simplicity, although it is an attractive topology because of the expected significant contribution in cross section. In the proceeding work, we will adhere to the historical standard of SUSY terminology: ‘gluino’ (\tilde{g}), ‘squark’ (\tilde{q}), and LSP ($\tilde{\chi}^0$). However, the results presented with simplified models are applicable to any new physics model that results in the same final state.

It is assumed that new particles are produced in pairs whose decay chains end in a stable weakly interacting massive particle described in Supersymmetry

¹ <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit>

² A joint project between the LHC experiments and the ROOT team to provide statistical tools built on top of a previously existing data modeling toolkit distributed through ROOT.

as the “lightest supersymmetric particle”, or LSP, as a consequence of R -parity conservation (see Chapter 1 for details).

The final states of T1 and T2 are different as seen in Figure 5.1. The left figure represents the T1 topology and it is clear from the diagram that the final state contains two quarks, two antiquarks, and two LSPs. This means that this topology will contain events with 4 jets (one for each quark or antiquark) and a source of E_T^{miss} . The right panel of this figure shows the T2 topology physical representation where there will be only two jets and a source of E_T^{miss} . The interpretation of what this will mean in regards to the efficiency of our search is presented later in this chapter.

Another topology of interest is the right panel of Figure 5.2, the T5ZZ topology. Much like T1, it begins with gluino pair production and its final state includes four jets (from quark and antiquark production); however, in addition it has a secondary cascade to LSP candidates and Z bosons. The SMS that include long cascade decays are interesting to study, since the amount of energy available for the LSP is less, for the same mother mass, as when there is a direct decay like T1. Comparisons of results using T1 and T5ZZ indicate the sensitivity of the search to a certain topology depends on details of the decays of produced SUSY particles.

The final set of topologies shown in Figure 5.3 are those including final states with third-generation quarks, t and b . These are of the same base gluino pair production as the T1 topology, but the secondary quarks from cascades of a gluino are specified to be the heaviest of all the quarks. This is significant because the masses for the u , d , s , c , b , t quarks are approximately 2.4 MeV, 4.8 MeV, 104 MeV, 1.3 GeV, 4.2 GeV, and 173 GeV respectively. Clearly there is a large difference in quark masses that would lead to very different final state topologies.³ This also provides

³ The top quark decays to 3 jets (via $t \rightarrow b + W \rightarrow b + q + q$ or $t \rightarrow b + W \rightarrow b + l + \nu$), which leads to different conditions necessary for jets (and leptons) too.

Table 5.1: Summary table of first-generation simplified models.

Prod.					
Name	Mode	Decay	Mass of LSP	Visibility	Gen.
T1	$\tilde{g}\text{-}\tilde{g}$	$\tilde{g}\rightarrow qq\tilde{\chi}^0$	massive	hadronic	1st
T2	$\tilde{q}\text{-}\tilde{q}$	$\tilde{q}\rightarrow q\tilde{\chi}^0$	massive	hadronic	1st
T5ZZ	$\tilde{g}\text{-}\tilde{g}$	$\tilde{g}\rightarrow qqZ\tilde{\chi}^0$	zero, massive	hadronic, di-leptons	1st

the ability to test how the selection process (of events described in Chapter 4) results in different efficiencies for each topology.

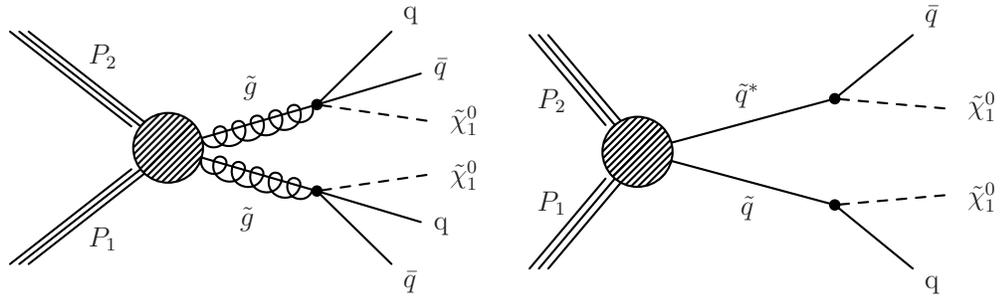


Figure 5.1: Diagrams of gluino pair production (T1, left) and squark pair production (T2, right).

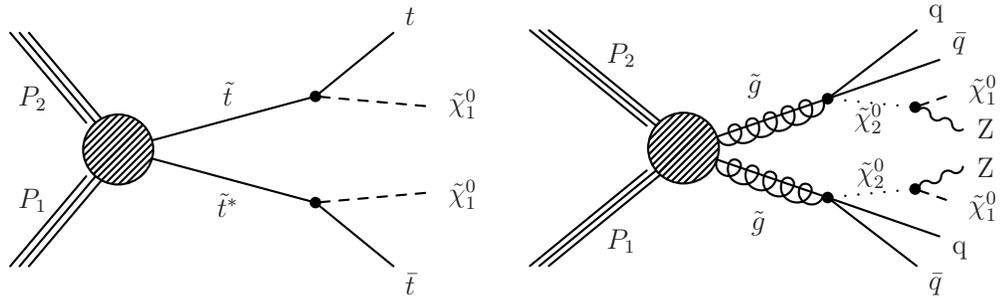


Figure 5.2: Diagrams of stop pair production (T2tt, left) and gluino pair production with intermediate Z boson (T5ZZ, right).

5.3 Interpretation with Simplified Models

In order to determine the cross section upper limit for these simplified model topologies, we need to know the fraction of the produced signal events that pass the event selection in the search intended to remove the majority of the background and

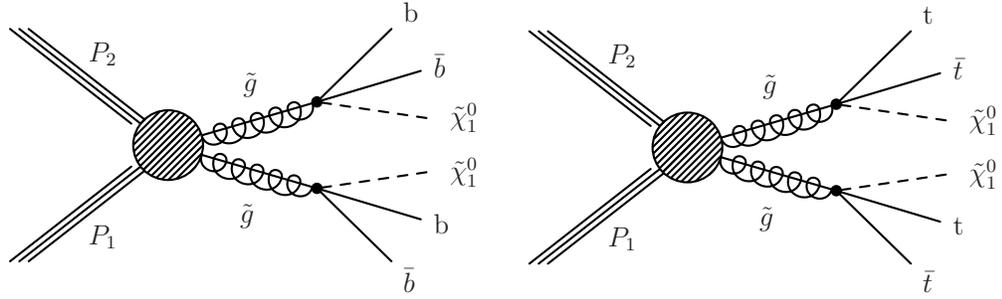


Figure 5.3: Diagrams of gluino pair production to bottom quarks (T1bbbb, left) and top quarks (T1tttt, right).

Table 5.2: Summary table of third-generation simplified models.

Name	Prod. Mode	Decay	Mass of LSP	Visibility	Gen.
T1bbbb	$\tilde{g}-\tilde{g}$	$\tilde{g} \rightarrow b\bar{b}\tilde{\chi}_1^0$	massive	hadronic	3rd
T1tttt	$\tilde{g}-\tilde{g}$	$\tilde{g} \rightarrow t\bar{t}\tilde{\chi}_1^0$	massive	hadronic	3rd
T2tt	$\tilde{t}-\tilde{t}$	$\tilde{t} \rightarrow t\tilde{\chi}_1^0$	massive	single lepton, di-leptons hadronic, single lepton	3rd

improve the signal/background ratio. This fraction is called the signal acceptance and efficiency. This is discussed below, followed by the results in the context of simplified models.

5.3.1 Acceptance Times Efficiency for Signal

The total signal efficiency of the search selection is shown within the simplified model space for gluino and squark pair production, as a function of the gluino or squark mass and the LSP mass. The M_{mother} vs $M_{\tilde{\chi}^0}$ distributions only display values in the lower half of the diagonal plane because the LSP must be lighter than that of the mother particle (due to basic kinematics). The signal efficiency increases for higher gluino and squark masses, and is low toward the diagonal region, where the mass splitting (of $M_{\text{mother}}-M_{\tilde{\chi}^0}$) is small and jets are produced with lower transverse momentum. The overall selection efficiency tends decrease for a fixed mother and

daughter masses as the number of patrons in the final state increases. Topologies like T2 have a higher efficiency because it contains only 2 jets, where T1 and T1bbbb (and T5ZZ) have lower efficiencies due to their having 4 (and 4–8) jets each.

In the limit calculation, the sources of uncertainty that are incorporated for each scan point include the jet energy scale and resolution, the lepton veto, the cleaning including the veto on large energy loss in masked ECAL cells, the trigger, the parton distribution functions (PDF), the luminosity, and the statistical uncertainty. Renormalization and factorization scale uncertainties do not apply here because they only influence the normalization of the reference cross section. It should also be noted that the presence of signal events in the background sample is negligible and not considered.

Each topology’s 95% CL upper limits, signal efficiency, and its uncertainty due to initial- and final-state radiation (ISR/FSR) is shown in Figures 5.4–5.10. The acceptance times efficiency from 14 search regions is evaluated for each of 14 search regions, and are combined together. The uncertainties are also evaluated for each of the search regions, and combined together by taking the acceptance-weighted average. Other plots such as the signal efficiency uncertainties due to parton distribution functions (PDF), jet energy scale (JES), and jet energy resolution (JER), are presented in Appendix A.

Initial- and final-state radiation comes from QCD effects (i.e. from colored particles). Radiated gluons can create an extra jet, which affects the number of jets in an event passing the selection. These extra jets may also greatly affect the event’s H_T and H_T^{miss} measurements, and therefore also change the efficiency for the signal. Thus, the signal efficiency becomes increasingly sensitive to the ISR modeling. In the diagonal region where the $M_{\text{mother}} - M_{\tilde{\chi}^0}$ difference is small, we expect only a small number of energetic jets from SUSY particle decays so the ISR jets play a more significant role. ISR modeling is not well understood, so we do not quote the

results near the diagonal region where the signal acceptance uncertainty due to ISR modeling can reach values larger than 50%.

Jet calibration is taken into account by correcting reconstructed jets to the particle-level. This calibration provides a corrected energy scale for reconstructed particles, grouped together by the jet algorithm. There are a number of effects that reduce the accuracy of the jet energy measurement. For instance, particles punch through the calorimeter without leaving all their energy, or they hit parts of the device which are malfunctioning. The part of the calorimeter measuring the jet may be hit by additional particles not belonging to the jet originated from the quark, affecting our measurement. From a rather complicated method using both real and simulated events, we are able to determine a correction to the particle energy/momentum measured in the detector. The correction is able to rescale the measured jet energy to the true jet energy, which essentially matches the energy of the originating quark (see Section 1.3). This is an essential and complicated process that can change the number of jets above the selection threshold, and H_T and H_T^{miss} values, thus affecting the event selection efficiency.

Jet energy resolution and linearity are key factors in separating signal events from backgrounds. The jet resolution differences between data and MC events vary with rapidity and the corrections are performed independently in five rapidity regions.

In order to predict the results of any hard-scattering process where hadrons are initial state particles it is necessary to know the parton distributions within the hadrons. The probability density for finding a particle with a certain longitudinal momentum fraction x at momentum transfer Q^2 is called a parton distribution function (PDF). The main reason why PDFs are important at the LHC is that at a hadron collider a detailed understanding of PDFs is necessary in order to obtain accurate predictions for any process. The determination of the PDF is based on the

“global fit” of various experimental measurements, in particular the deep inelastic scattering cross section measurements, which is beyond the scope of this thesis [32].

5.3.2 Results

The measured cross section upper limits are compared to a typical reference next-to-leading-order cross section from PROSPINO [33] (a program for the production of supersymmetric particles in next-to-leading-order QCD). For squark pair production (T2) the reference cross section refers to the squark-antiquark cross section with four light flavors included, which is used to convert upper limits on the production cross section to limits on new particle masses. For the topologies T1, T1bbbb, T1tttt, T5ZZ, the reference cross section refers simply to that of the gluino-gluino pair production.

The 95% CL upper limits on the cross sections for each of the topologies are displayed as a function of the mother and LSP mass. Several contours are visible on these plots determined by our sensitivity to each topology. These contours illustrate where the reference cross section (three times the reference cross section and one-third times the reference cross section) and the upper limit on the cross section intersect. The choice of the three or one-third factor is arbitrary, but chosen to provide uncertainties due to changes in the reference cross section occurring if a final state has a branching ratio different than 1.0, or if the produced particle has degenerate states or a different spin than assumed. The 95% confidence level of the upper limit of the signal cross section is compared to a reference cross section (e.g. gluino, squark, stop pair production) that is specific to each topology. The region underneath a curve where the upper limit on the signal cross section and the reference cross section intersect are excluded from the possible mass combinations for M_{mother} and $M_{\tilde{\chi}^0}$. Generally speaking, the lower the mass, the higher the production cross section, thus the exclusion of the lower mass region.

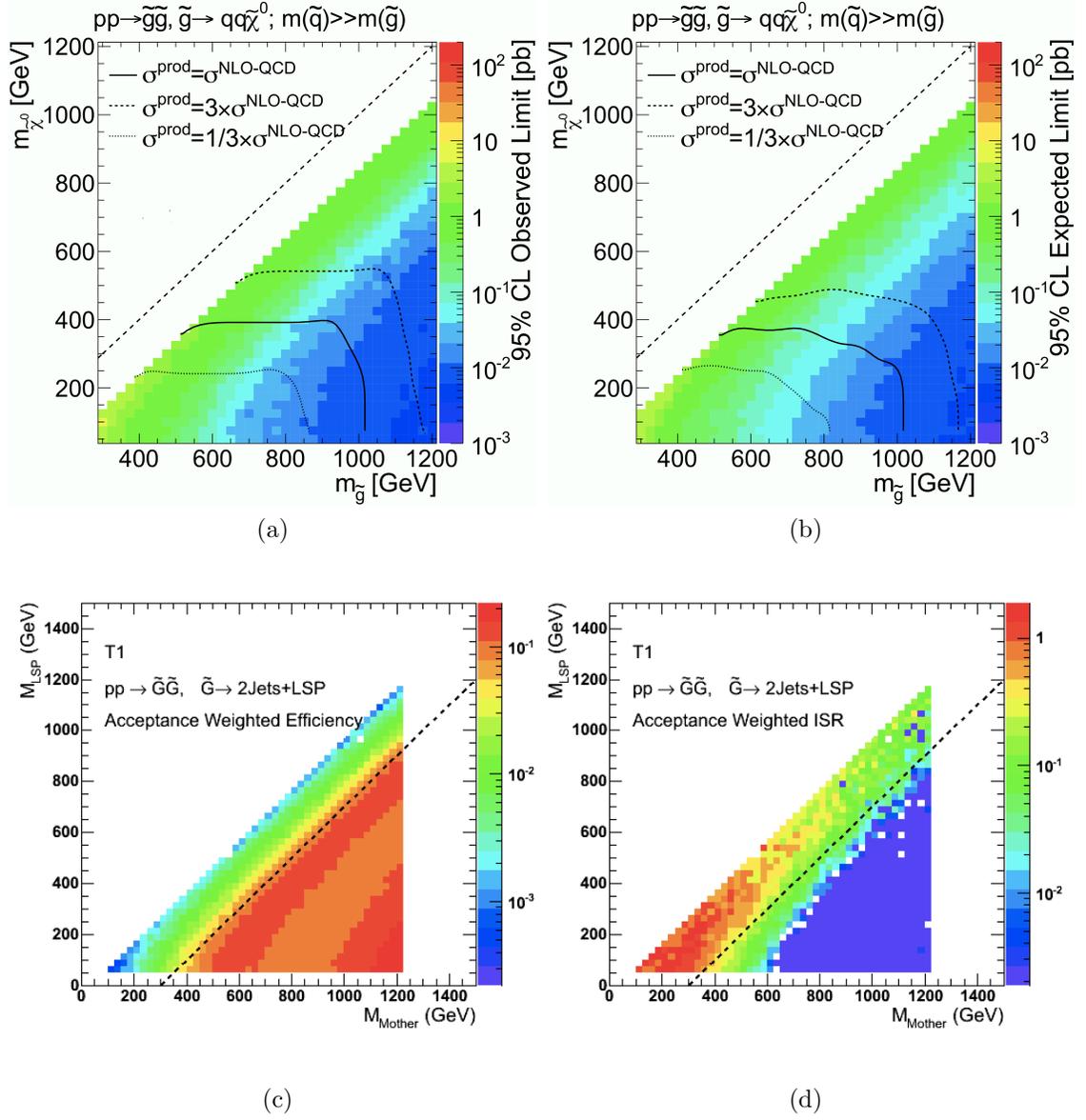


Figure 5.4: T1 (a) The observed and (b) expected 95% CL upper limits on the $\tilde{g}\tilde{g}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.

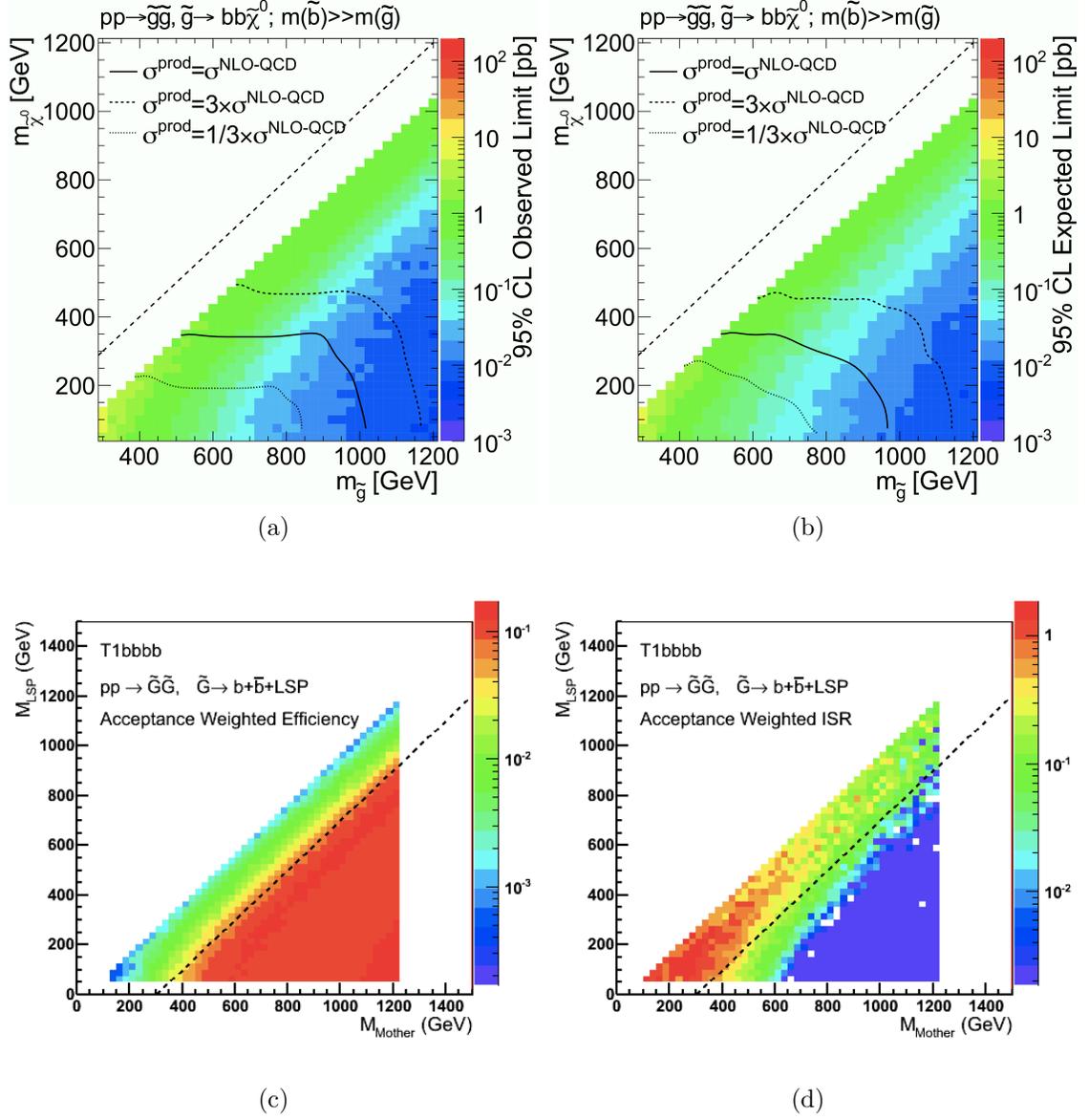


Figure 5.5: T1bbbb (a) The observed and (b) expected 95% CL upper limits on the $\tilde{g}\tilde{g}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.

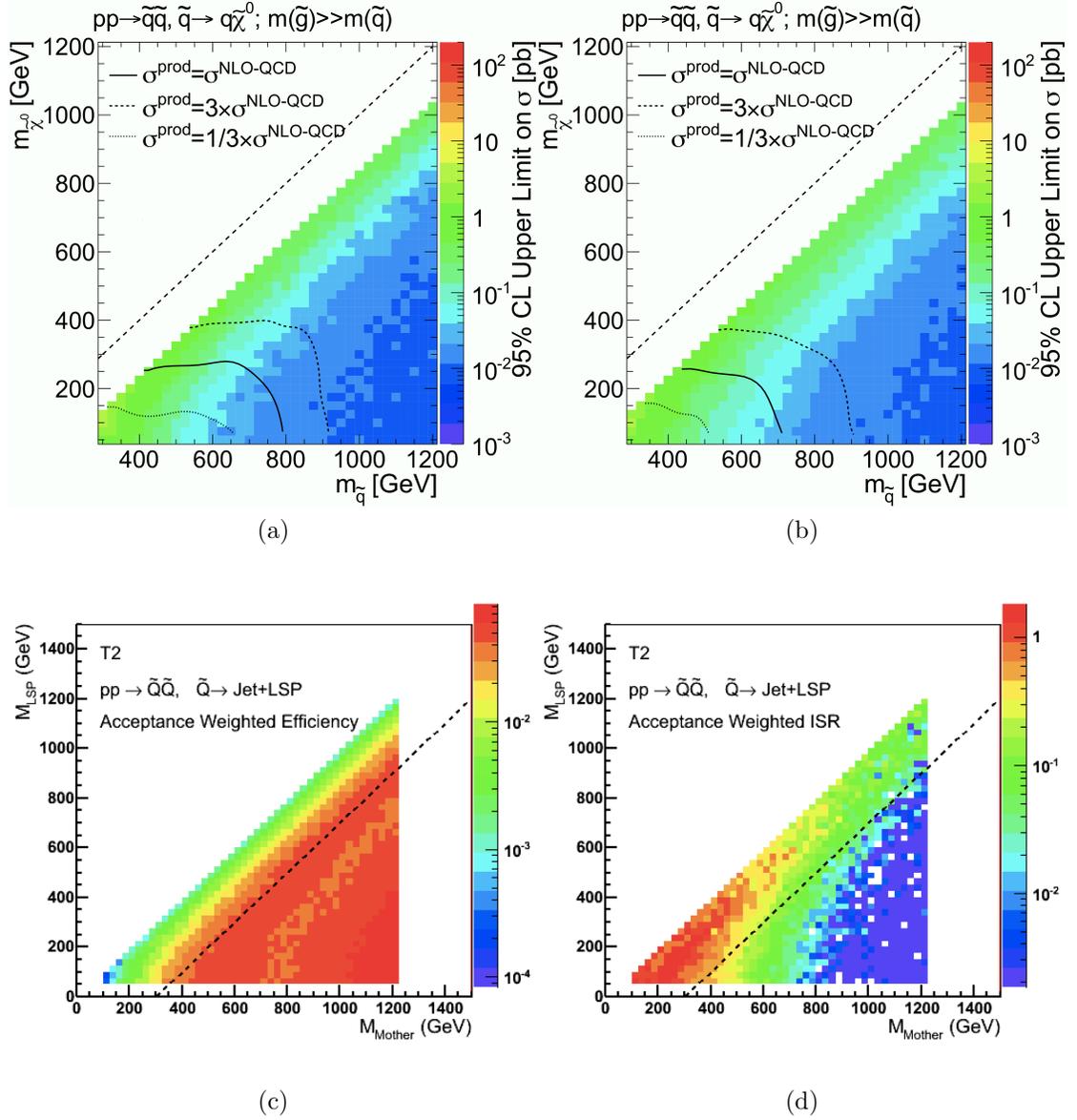


Figure 5.6: T2 (a) The observed and (b) expected 95% CL upper limits on the $\tilde{q}\tilde{q}$ cross sections in the $m_{\tilde{q}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.

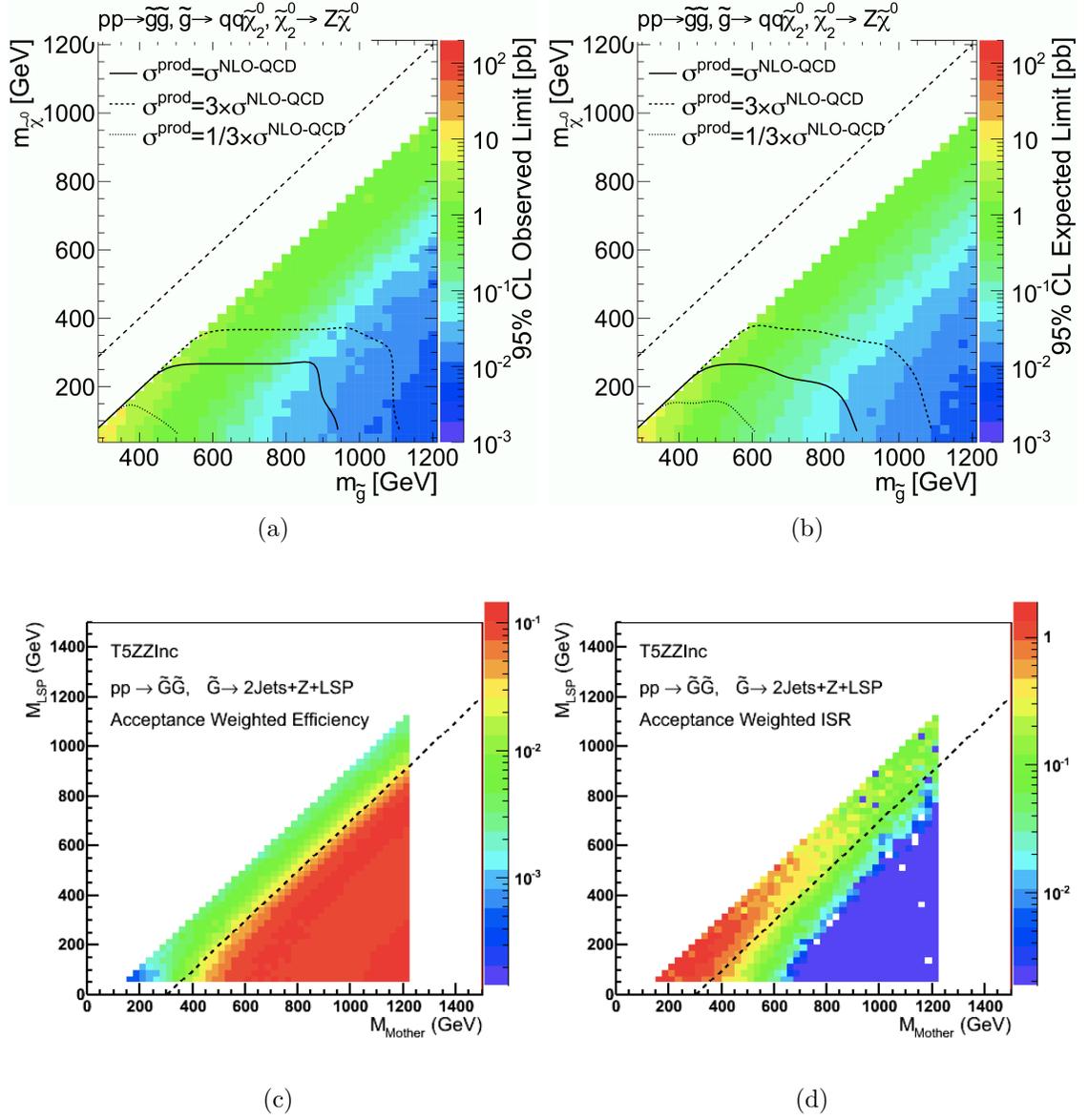


Figure 5.7: T5ZZ (a) The observed and (b) expected 95% CL upper limits on the $\tilde{g}\tilde{g}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.

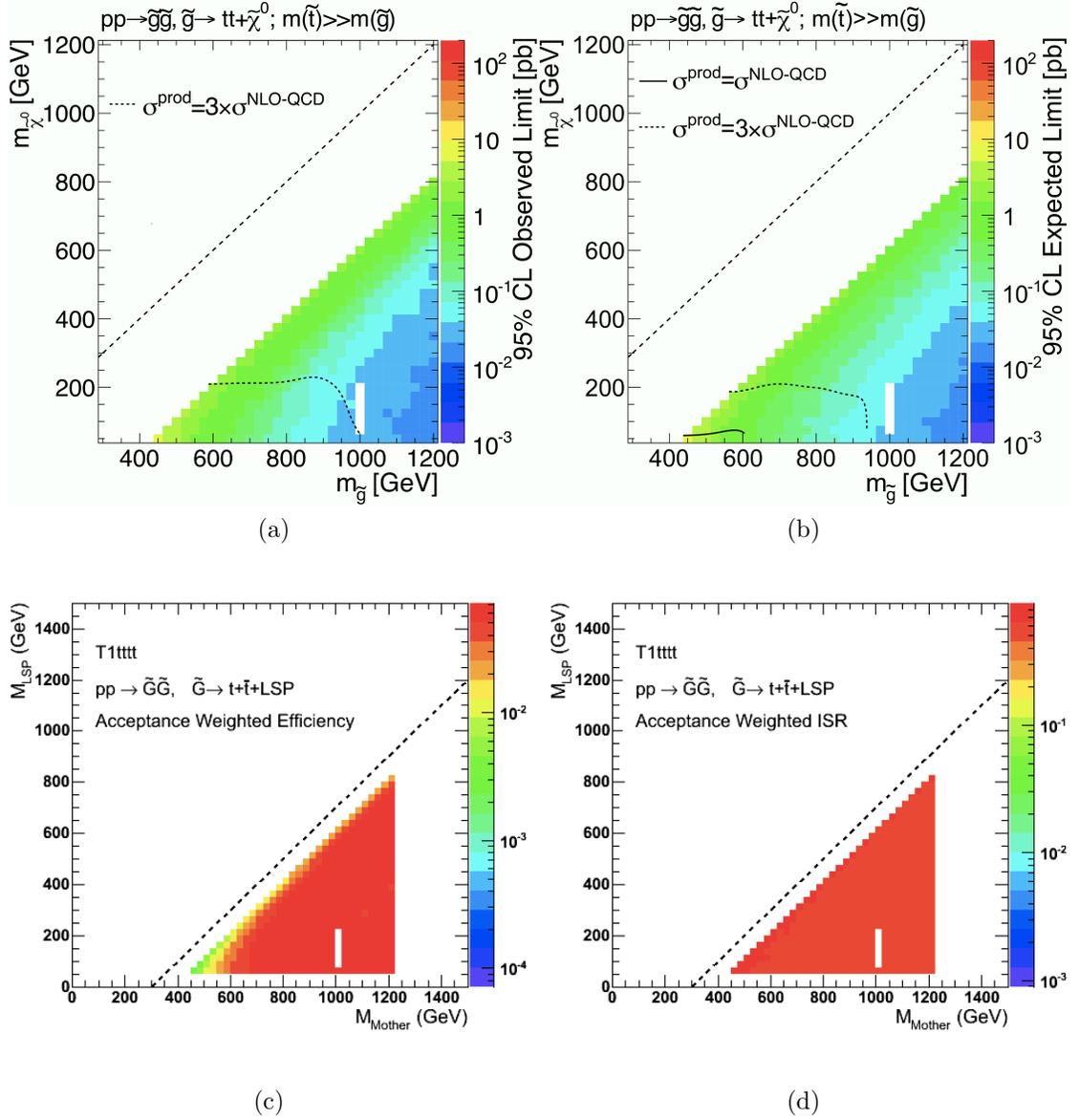


Figure 5.8: T1tttt (a) The observed and (b) expected 95% CL upper limits on the $\tilde{g}\tilde{g}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.

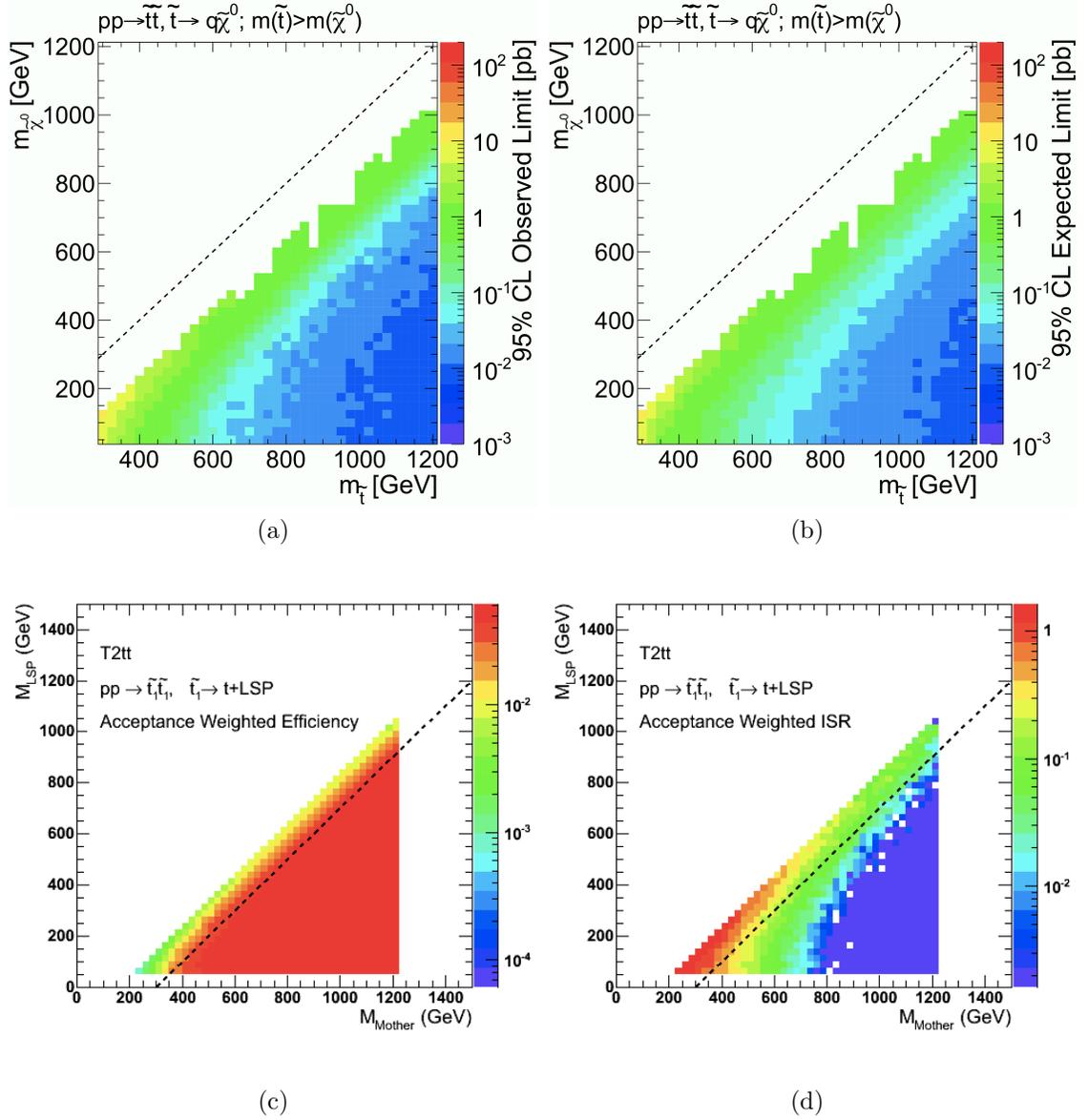
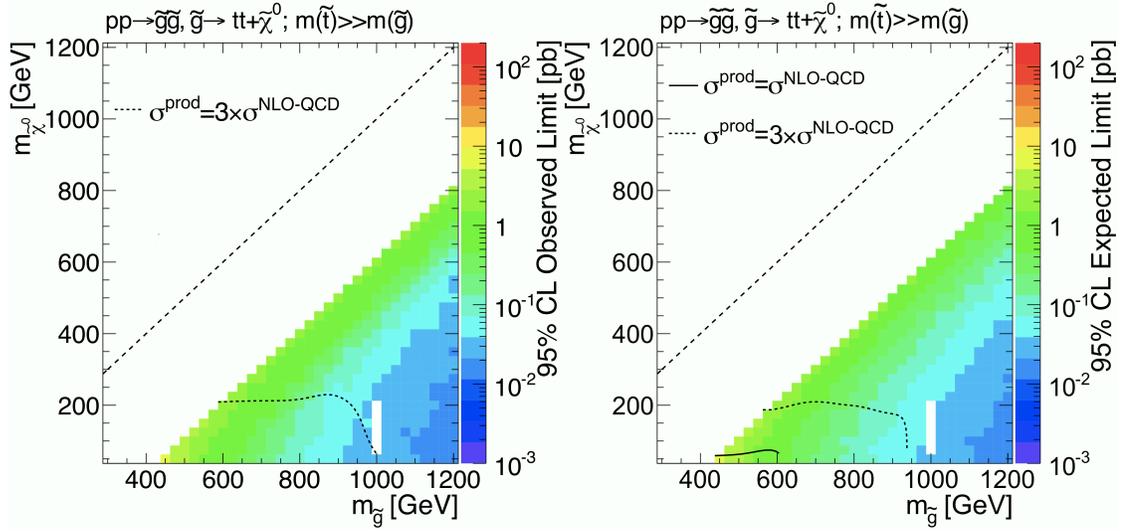
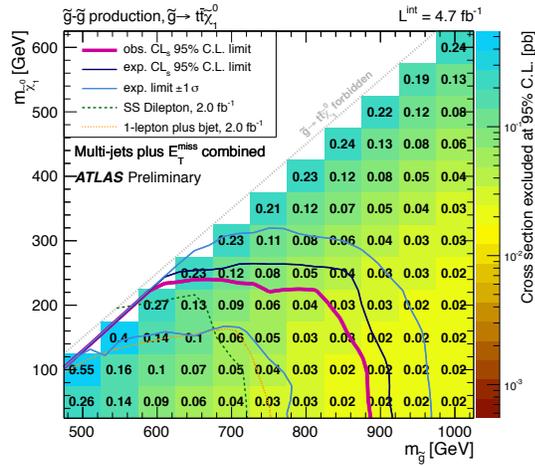


Figure 5.9: T2tt (a) The observed and (b) expected 95% CL upper limits on the $\tilde{q}\tilde{q}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. (c) Signal selection acceptance times efficiency for all 14 search regions combined. (d) Signal selection acceptance times efficiency uncertainty due to ISR modeling for all 14 search regions combined.



(a)

(b)



(c)

Figure 5.10: T1tttt (a) The observed and (b) expected 95% CL upper limits on the $\tilde{g}\tilde{g}$ cross sections in the $m_{\tilde{g}}$ and $m_{\tilde{\chi}^0}$ and planes. The contours where the single cross section, three times this cross section, and one third times this cross section can be excluded are also shown. Bottom: ATLAS public results of observed and expected 95% CL upper limits on the same plane.

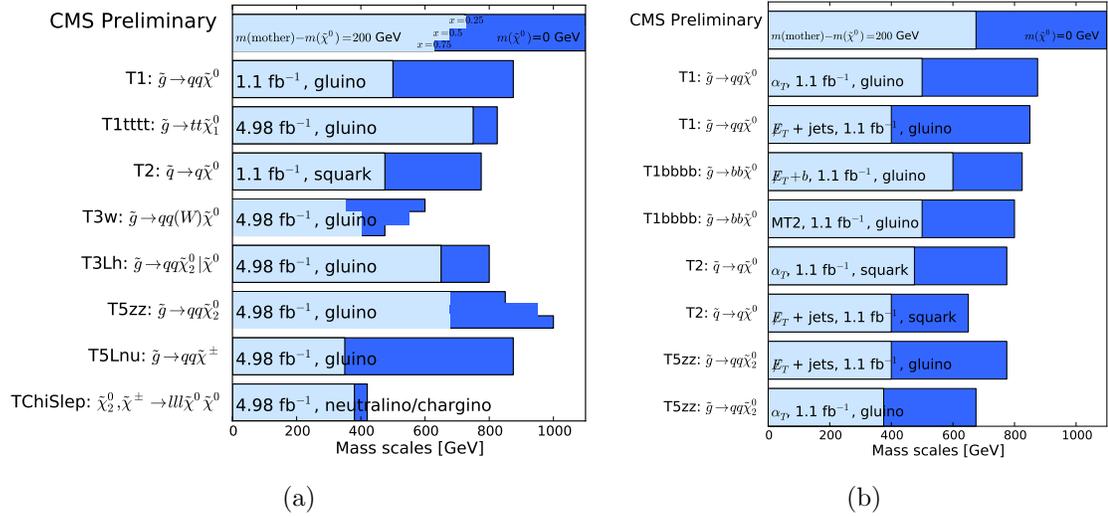


Figure 5.11: (Left) Best exclusion limits for gluino and squark masses, for $M_{\text{mother}} - M_{\tilde{\chi}^0} = 200$ GeV (light blue) and $M_{\tilde{\chi}^0} = 0$ GeV (dark blue), for each topology, for all results. (Right) Exclusion limits for gluino and squark masses, for $M_{\text{mother}} - M_{\tilde{\chi}^0} = 200$ GeV (light blue) and $M_{\tilde{\chi}^0} = 0$ GeV (dark blue), for each topology, for hadronic results.

Topologies T1 and T1bbbb show nearly identical 95% limits. We naively expect this to be the case because our analysis group does no b jet tagging and removal from our samples. The way our cuts and selections are chosen, T1 should (and does) have the best exclusion curves of all the topologies reviewed here due to it having more partons in the final state and our analysis requiring a minimum of 3 high- p_T jets. T2 has a slightly reduced exclusion contour than that of T1, as anticipated, because of the larger expected cross section for a color octet particle than for a color triplet particle (see details in the QCD section of Chapter 1).

T5ZZ suffers from a loss in exclusion compared to T1 primarily due to its long cascade and having 6–8 jets as described above, but is still an important model to investigate because it demonstrates how the limits and sensitivity to the SUSY signal depend on the decay kinematics. T1tttt shows only a few of the exclusion curves due in part to the fact that we have much smaller statistics for the T1tttt sample than for T1. We expect this reduction in signal acceptance times efficiency

because the top quark decays to a b quark and a W boson, and the W boson then decays either two jets or lepton + ν . If W decays to lepton + ν , this event is rejected. Our jets + E_T^{miss} search rejects such events because of the lepton veto that is necessary to suppress the W and top quark production background. The signal acceptance time efficiency for one of the most sensitive regions ($H_T > 1400$ GeV and $H_T^{\text{miss}} > 200$ GeV) on T1tttt is $\sim 1/3$ of that for T5ZZ. We have much less signal acceptance times efficiency of T1tttt in high H_T^{miss} bins than for T5ZZ due to lepton veto.

The obtained 95% CL limits for T2tt are included, but due to insufficient sensitivity of this particular topology no 95% upper limit on the cross section can be drawn. The T2tt 95% CL upper limit results also show an unusual diagonal edge that the other topologies do not have. The jagged edge shown is due in part to not having sufficient input information (the edge of the diagonal region in general suffers from high ISR uncertainties) and suffers from convergence failures in the limit calculation.

From the previous discussion, it is worth noting that the limit on the gluino or squark depends strongly on the LSP mass. The 95% CL limit is significantly weaker when $M_{\text{mother}} - M_{\tilde{\chi}^0}$ is small, due to there not being enough energy available for multiple jet production. Currently, we do not find evidence of SUSY, but it may be because the SUSY realized in nature is in the phase space which is hard to access.

Figure 5.10 is an addition to Figure 5.4 in that it compares the results from the CMS jets + E_T^{miss} search to those from ATLAS in the topology T1tttt [34]. Unfortunately, at this point ATLAS shows a larger exclusion curve than our results, indicating that their search is more sensitive to this particular topology. This is not surprising because ATLAS specifically optimized their search by looking at very high jet multiplicity, which is what CMS is currently trying to do as well with some improvements.

In regards to exclusion results from all CMS search methods and on different topologies, Figure 5.11 shows a summary of all publicly released results. The left figure puts together all the best exclusion limits from both leptonic and hadronic searches. The luminosity of the samples used for each of the topologies is listed for each and whether the topology focuses on squark or gluino masses. The figure on the right shows the exclusion limits of only the hadronic results (similar to results shown in this paper). We include this figure to show that other groups are focusing on different key variables other than E_T^{miss} (e.g. α_T), or requiring a jet identified as a b -quark jet (e.g. $E_T^{\text{miss}} + b$ search). The $E_T^{\text{miss}} + b$ search is expected to show more sensitivity to the signal enriched in b 's. The different analyses performed among CMS groups have complementarity, but also differing sensitivities to the various signal topologies. Hopefully, this will help to find any SUSY signal that may appear in the near future.

CHAPTER SIX

Conclusion

A search for Supersymmetry was performed that focused on the key observable, missing transverse energy (E_T^{miss}). The significance of this observable cannot be overstated, as it is one of the primary experimental signatures that is predicted by theory and indicates the signal of the lightest stable supersymmetric particles. These lightest supersymmetric particles are also candidates for constituents of the dark matter observed in astronomical observations. Attempts to improve measurement of E_T^{miss} are of paramount importance for SUSY events to be discovered.

In 2011 the LHC collided two bunches of protons every 50 ns and approximately 10 proton-proton collisions occurred in each bunch crossing. This number is increasing over time, in 2012 the average is more than 20. This increase improves the amount of the data we can collect in a given time; however, it tends to deteriorate the resolution of the missing E_T measurement. Reported in Chapter 3 was a study to improve the missing E_T resolution using the time information from a component of the Compact Muon Solenoid (CMS) particle detector called the electromagnetic calorimeter. The study on the potential impact of applying the timing cut in the electromagnetic calorimeter end-cap (EE) for particle flow reconstruction showed a clear reduction in out-of-time pileup effects. It was found that by removing the reconstructed hits with the `kOutOfTime` flag in the EE with a 2 GeV energy threshold, the E_T^{miss} resolution improves by 1–2 GeV in quadrature at ~ 10 out-of-time collisions. It is known that both in-time and out-of-time pileup interactions degrade the E_T^{miss} measurement, and we provided a means for decreasing these effects using rehit timing information.

Search specifications centered around the presence of jet + E_T^{miss} in possible SUSY events were laid out in Chapter 4. By beginning with a search through CMS 2012 data for high E_T^{miss} events, a more in-depth explanation of the type of events that are of interest to the overall SUSY search was described. Additionally, the types and contributions of background standard model processes were discussed and tabulated to provide input parameters for the confidence limits presented in Chapter 5.

Traditionally, the results of SUSY-inspired searches by CMS have been interpreted in terms of the Constrained Minimal Supersymmetry extension of the Standard Model (CMSSM). It predicts a rich spectrum of experimental signatures, but with built-in relations between particle properties that affect the rate of signal events and kinematics of the final state particles. The Simplified Model Spectra (SMS), unlike CMSSM, assume that a particular decay signature can be realized without specifying the exact mechanism, offering the possibility to overcome small branching ratios.

Chapter 5 provides a description of the topologies from simplified models and their sensitivities with search parameters that were setup in the previous chapter. Uncertainties for the acceptance times efficiency (of the event selection on the signal) such as those encountered by initial- and final-state radiation, jet energy scaling, jet energy resolution, and parton distribution functions for each of the topologies were presented for each topology (and combined from each of the 14 H_T and H_T^{miss} regions).

The 95% confidence level limits on the cross section for various topologies were explored, where the M_{mother} vs $M_{\tilde{\chi}^0}$ distributions only display values in the lower half of the diagonal plane. This is due to the requirement that the LSP be lighter than that of the mother particle (basic kinematic principles). Results for T2 and T5ZZ produced slightly reduced 95% CL than those of T1 and T1bbbb, but these

results were anticipated for reasons described in the main body of the text. $T1tttt$ suffered from reduced acceptance times efficiency due to branching ratios of the top quark, and showed further reduced 95% CL upper limits from even that of $T2$. $T2tt$ did not suffer from the lack of acceptance times efficiency that $T1tttt$ had, but the search parameters set by our analysis group were not sensitive to this particular topology and therefore had no 95% confidence level on the upper limit of the cross section.

Clearly much work is to be done for the search for Supersymmetry, but significant strides have been made in the last few years. Many beyond-the-standard-model theories exist, but only methodical searches from many experimental facilities will be able to determine which one will prevail.

APPENDICES

APPENDIX A

Additional Material for the Simplified Model Spectra Studies

This section includes many additional figures from Chapter 5 that contain details that are finer points of the analysis. Specifically, the uncertainty on the signal acceptance-weighted efficiency (fraction of the signal events that pass the event selection for the search) due to the jet energy scale (JES), jet energy resolution (JER), and parton distribution functions (PDF) for all topologies listed in Chapter 5 are included here. A brief description of ISR/FSR, JES, JER, and PDF are included in the main body of the text.

Additionally, Figure A.1 shows that each of the 14 search regions described in the text produce their own efficiencies and uncertainties. In order to obtain the combined uncertainties from the 14 exclusive search regions, each bin with its uncertainty (or efficiency) is multiplied by that bin’s acceptance, and they are added together and then normalized using the sum of the acceptances for all bins. In other words, the uncertainties in the 14 search regions are combined based on the acceptance of each region, reflecting the contributions of each search region to the search sensitivity.

Figures A.2–A.7 show a dashed diagonal line, which is the “cutoff” region for the topology. We don’t quote the results where the ISR uncertainty exceeds $\sim 50\%$ because the acceptance times efficiency evaluation becomes unreliable due to the strong dependence on the ISR modeling. The uncertainties generally tend to increase toward the diagonal region where the mass splitting (of $M_{mother}-M_{\tilde{\chi}^0}$) is small and jets are produced with lower transverse momentum, so that the signal selection efficiency is more sensitive to e.g. jet p_T threshold and jet energy scale and resolution uncertainty.

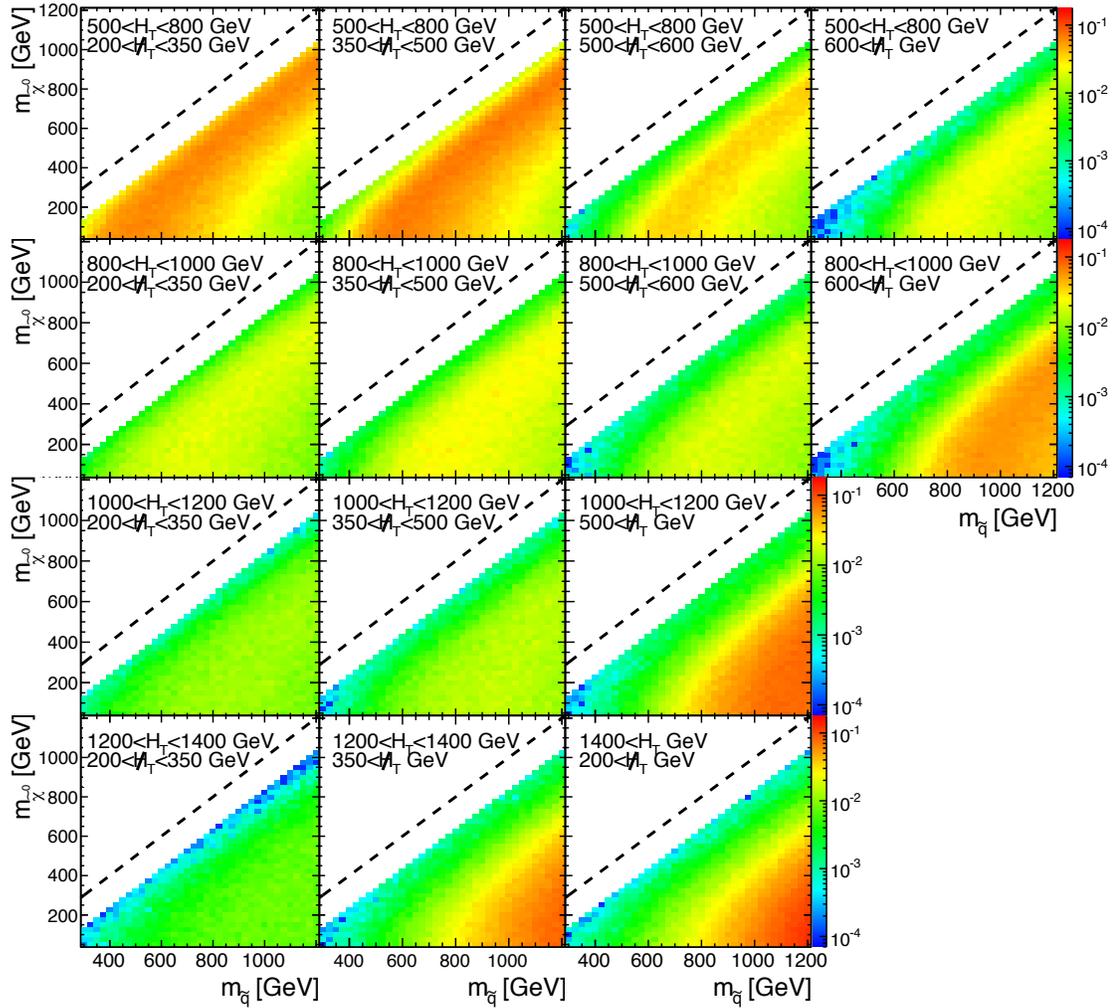
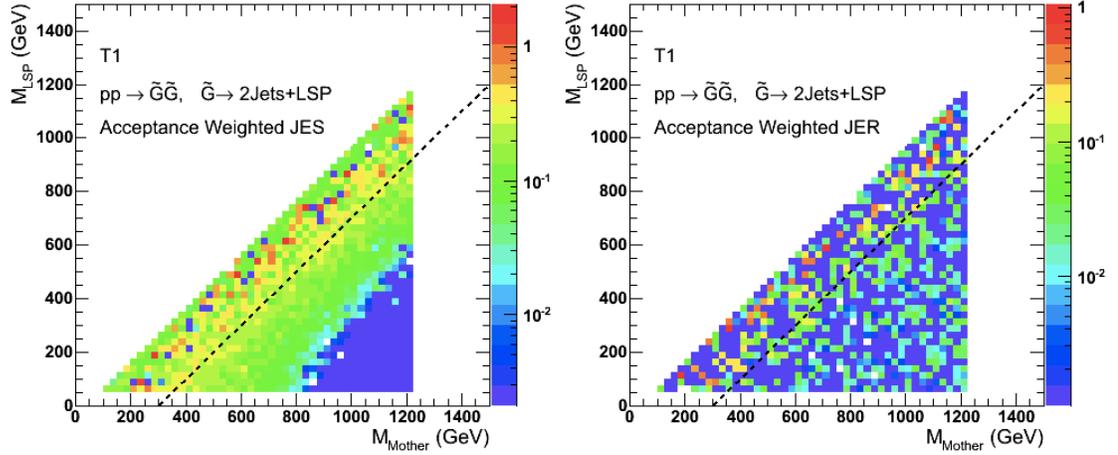
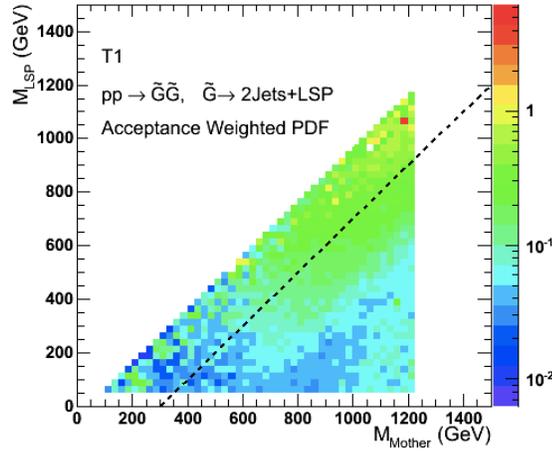


Figure A.1: T1 acceptance \times signal efficiency for each of the 14 bins, listed horizontally for constant H_T and varying H_T^{miss} regions and vertically for varying H_T and constant H_T^{miss} .



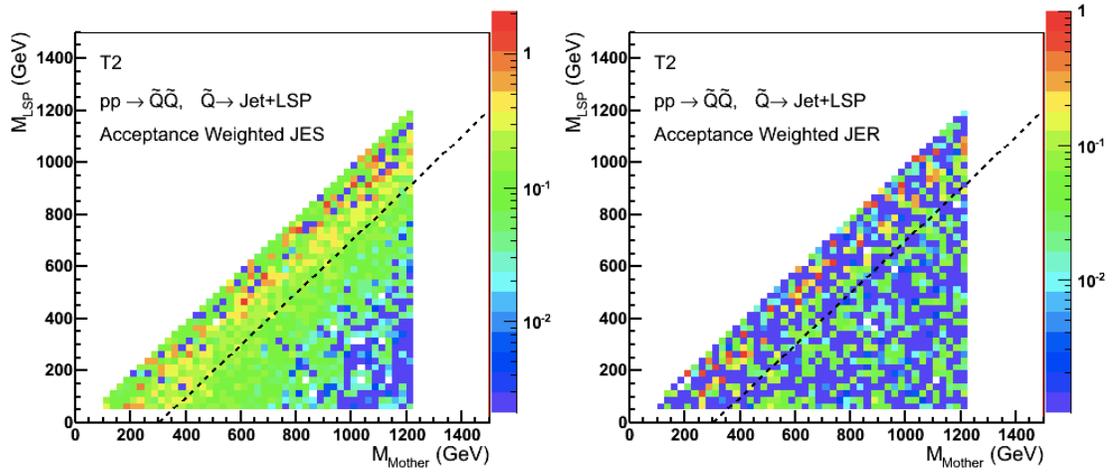
(a)

(b)



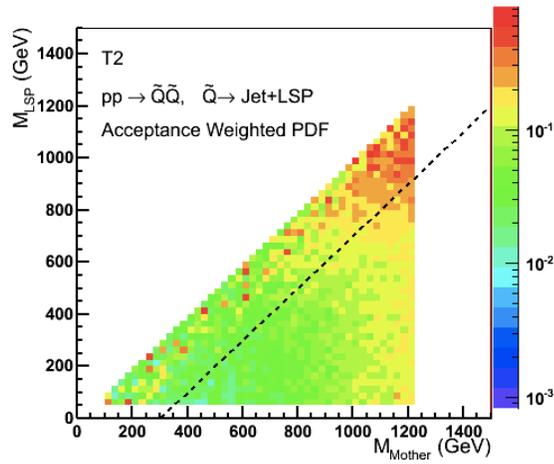
(c)

Figure A.2: SMS T1 topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.



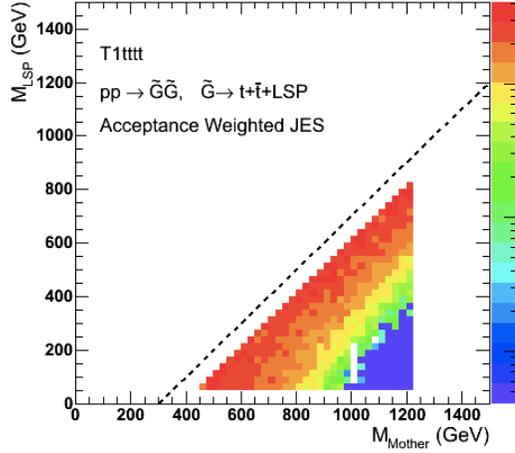
(a)

(b)

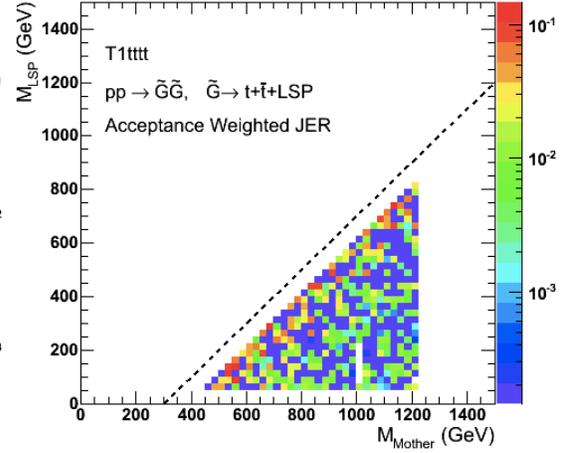


(c)

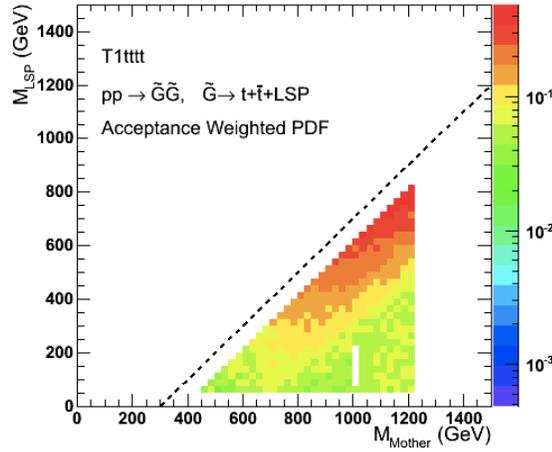
Figure A.3: SMS T2 topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.



(a)

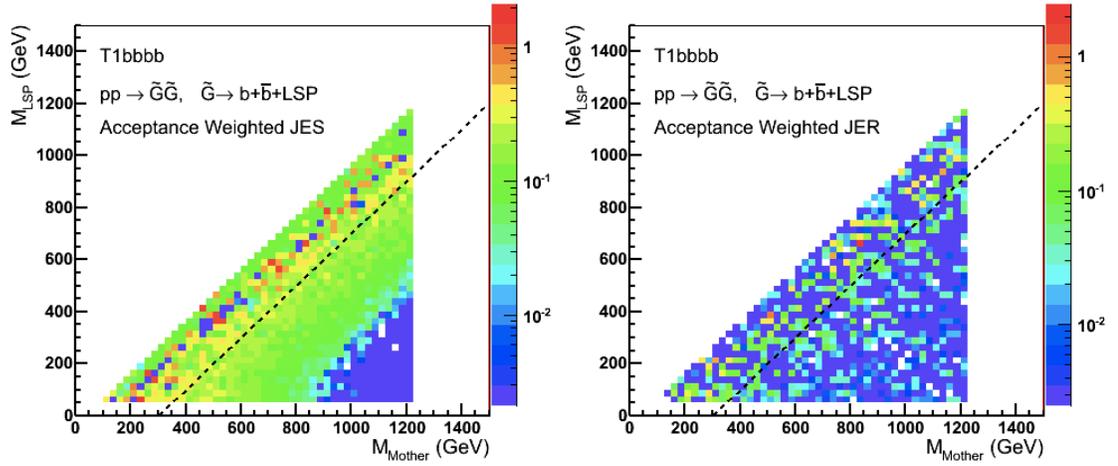


(b)



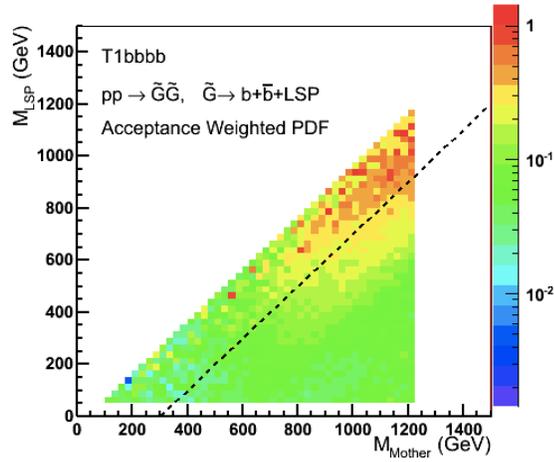
(c)

Figure A.4: SMS T1tttt topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.



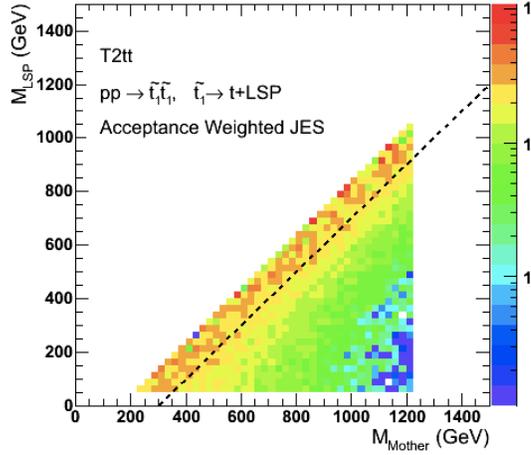
(a)

(b)

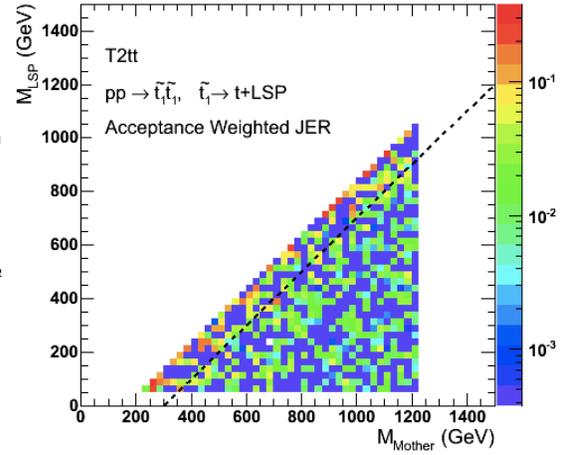


(c)

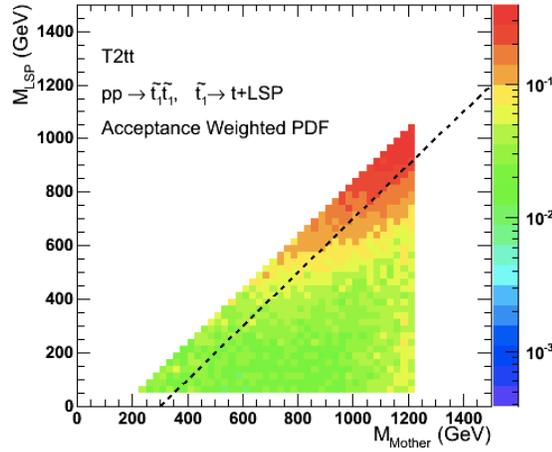
Figure A.5: SMS T1bbbb topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.



(a)

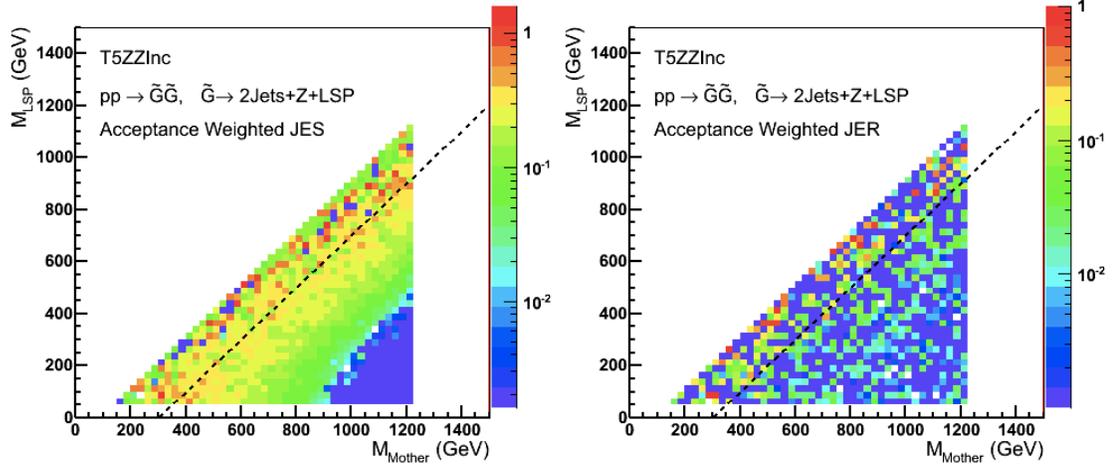


(b)



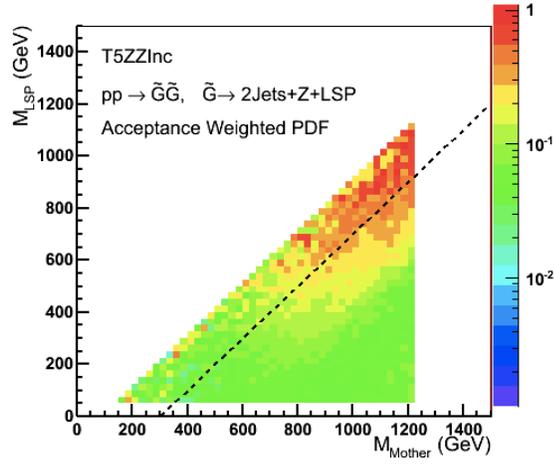
(c)

Figure A.6: SMS T2tt topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.



(a)

(b)



(c)

Figure A.7: SMS T5ZZ topology signal acceptance times efficiency uncertainty due to (a) JES, (b) JER, and (c) PDF from the 14 search regions combined.

APPENDIX B

Hadronic Calorimeter Reconstruction Software Validation

A brief introduction to the CMS detector components was included in Chapter 2, which included a section regarding the hadronic calorimeter. A discussion on CMS software validation in relation to this device will follow.

B.1 Software Validation

Nearly every week, the CMS software stack comprised of over 2 million lines of code is updated and must be checked for consistency. In order to assure the quality of each release, a validation process was developed for the authors to compare the performance of a new release to a previous one as reference. Using reconstructed datasets of real and MC samples any inconsistencies can be detected and fixed in a timely fashion.

A *Pre-Release* consolidates the current state of the code base to test interdependencies between different software components that are being modified in parallel. Once the current software component developments are completed, a *Release* is distributed and installed on all levels of the CMS tiered computing structure. All new features must then be placed into the next release cycle.

I began doing HCAL data validations back in October 2010 with the CMSSW_3_9_0_pre5 release, finalizing my contributions to this service work with the CMSSW_6_0_0_pre5 release. Over one and a half years, over 80 different validations were performed, and results have ranges from a successful validation where there was no difference in the HCAL calorimetry towers (CaloTowers), reconstructed hits (rechits), or read out box noise (RBX noise), to differences encountered in many of these structures that were originally unexpected.

I contributed to the CMS internal website¹ where the section entitled “Running crab jobs for data RelVal” describes how to properly setup and use existing packages to validate one release with another. HCAL data validations that have been completed are posted locally to the Baylor HEP website.² All listings on the website are in the format “new release”_vs_“old release”_RelVal/. Occasionally there are multiple versions of the same release if differences were found in the components described previously, which are denoted by “V#”. Once validations are complete, the results must be communicated on the RelVal HyperNews forum, where other experts and developers can review the findings and determine any software changes that may have introduced a bug.

During validations, it is necessary to review the listed HCAL CMSSW Release Notes which describe any changes made between releases and any expected effects on the validation.³ The set of database tags, which together define the conditions of the data, are collected together in a “Global Tag”, which is itself stored in the database. This decouples the conditions of the database from the CMSSW release, since different Global Tags can be used with a given CMSSW release. When this occurs, differences seen in the validation may be due to changes in this Global Tag.⁴

Figure B.1 is one of the many plots reviewed to determine if the validations are successful. This clearly shows that the number of events compared in each sample are identical, and the histograms completely overlap with no difference in either dataset. However, there are many cases in which this doesn’t occur, which will be described next.

¹ <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HcalValidation>

² <http://hep.baylor.edu/cmshcal/Validation/>

³ <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HcalReleaseNotes>

⁴ <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGGuideFrontierConditions>

Figure B.2, for example, shows a comparison where there is a slight difference in the number of events for each sample, but a large difference in the time distribution, specifically at high negative times. Upon investigation, it was found that the GlobalTag was changed between these two datasets (shown in bold):

- /MinimumBias/CMSSW_4_4_1-**GR_R_44_V6**_RelVal_mb2011A-v1/RECO
- /MinimumBias/CMSSW_4_4_0-**GR_R_44_V5**_RelVal_mb2011A-v2/RECO

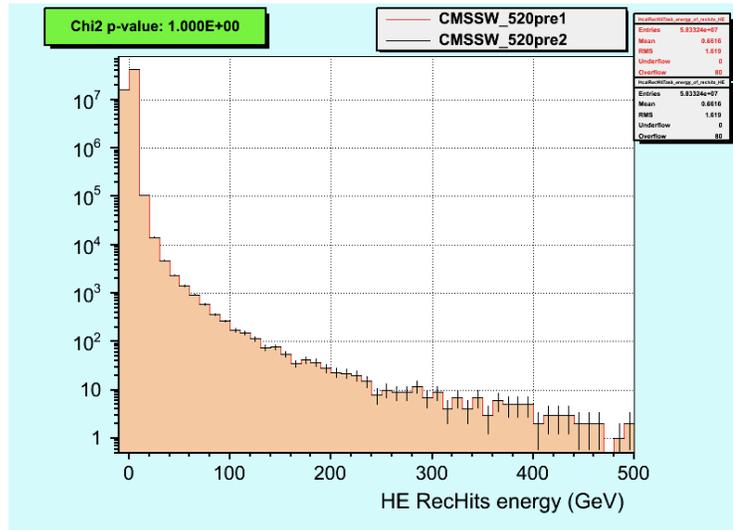


Figure B.1: Energy distribution of reconstructed hits for the HCAL end-cap region. Result from validation CMSSW_5_2_0_pre2 vs CMSSW_5_2_0_pre1 showing a successful validation, where no differences are found between the two releases.

There were no code changes during this validation that would account for any differences, however, this updated GlobalTag accounts for the differences that were found. The GlobalTag change is described as new HCAL time slices being initiated for the 4.4.1 dataset. Overall, the validation is successful because the changes seen are accounted for by the new HCAL signal integration time setting..

The next example (Figure B.3) is a validation in which the mean values of the HB (Calo) E_T^{miss} are different between releases. However, when the HB Calo-Towers ECAL and HCAL energy plots are investigated, it is found that the ECAL component changes while the HCAL does not. The differences found in this release

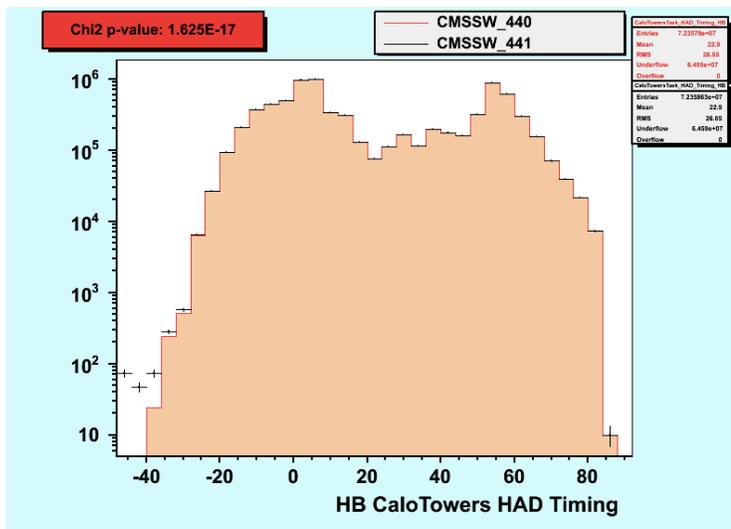


Figure B.2: Time distribution for the HCAL CaloTowers in the barrel region in ns. Result from validation CMSSW_4_4_1 vs CMSSW_4_4_0 showing that differences are found between the two releases in -50ns to -30ns region.

are thus due to ECAL changes that are not of interest to this validation. Overall, the validation is successful, but the differences are noted to inform developers that changes were seen in our analysis that can be double-checked with ECAL results. ECAL developers attribute this difference to defining new constants that convert ADC signals to GeV and new laser correction constants that are included in the CMSSW_4_2_0 release.

Another more recent example of time slice changes occurs during the CMSSW_5_0_0_pre7 release where the HF timing was changed from 2 time slices (TS) to 1 TS, and the HB and HE were both switched from 4 TS to 2 TS. The timing differences in Figure B.4 are much more apparent than the slight modification made in the CMSSW_4_4_1 release.

This task was my major “service work” for the CMS Collaboration. For this contribution, I have been a signing author of the all CMS publications since Fall 2011. The CMS collaboration has published 55 papers since that time.

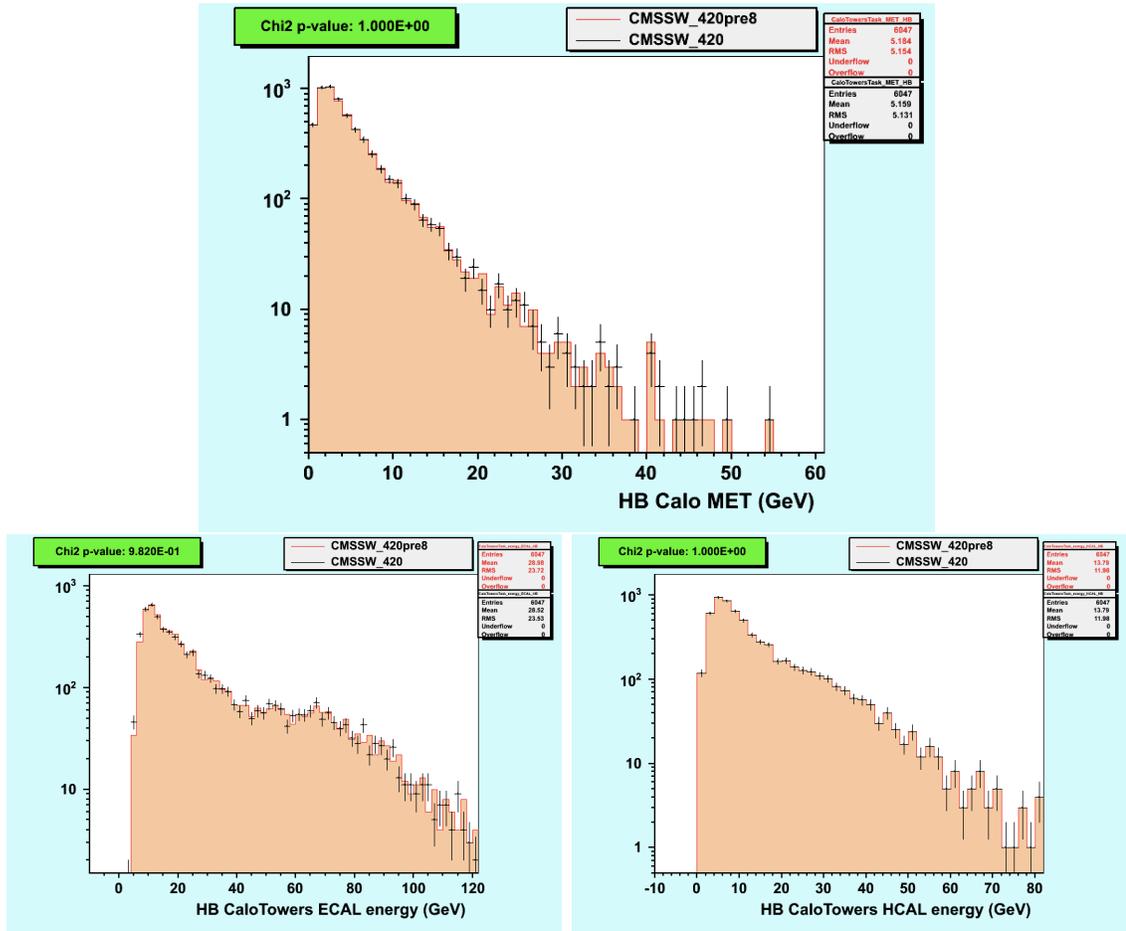


Figure B.3: Top: Barrel E_T^{miss} distribution with differences seen in the 20–60GeV range. Bottom: (Left) Barrel calorimeter towers ECAL energy distribution with differences seen in the 80–120 GeV range. (Right) Barrel calorimeter towers HCAL energy distribution. Result from validation CMSSW_4_2_0 vs CMSSW_4_2_0_pre8.

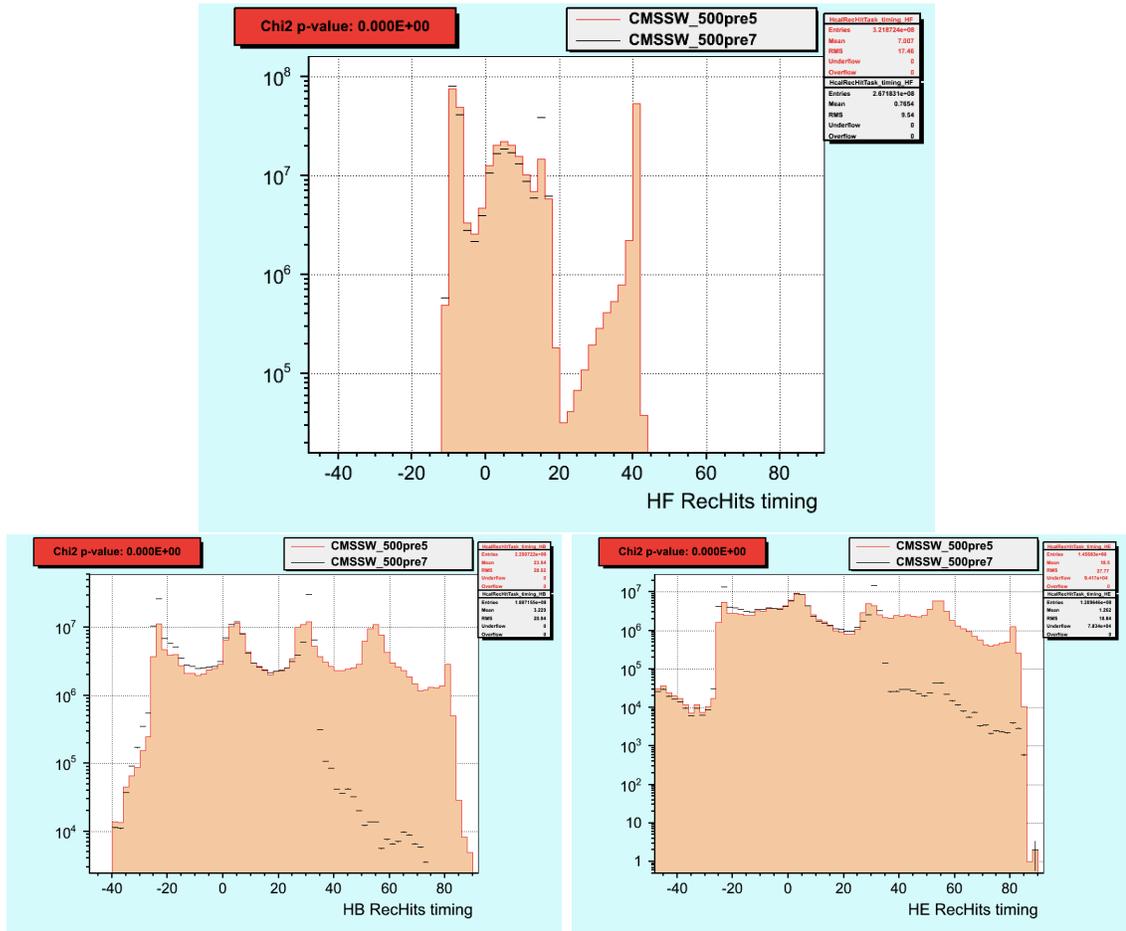


Figure B.4: Top: HCAL forward RecHits timing distribution [ns] with no 5_0_0_pre7 contribution in the 20–40ns range. Bottom: (Left) HCAL barrel RecHits timing distribution ns with difference in 20–80 ns range. (Right) HCAL end-cap RecHit timing distribution ns with differences found in 20–80 ns range. Result from validation CMSSW_5_0_0_pre7 vs CMSSW_5_0_0_pre5.

BIBLIOGRAPHY

- [1] K. Perkins, *Introduction to High Energy Physics*, 4th ed. (Cambridge University Press, 2000).
- [2] D. Griffiths, *Introduction to Elementary Particles*, 2nd ed. (Wiley-VCH, 2008).
- [3] John F. Donoghue, “Introduction to the Effective Field Theory Description of Gravity,” arXiv:gr-qc/9512024v1, 1995.
- [4] Francis Halzen & Alan D. Martin, *Quarks and Leptons: An Introduction Course in Modern Particle Physics*, (John Wiley & Sons, 1984).
- [5] Graham G. Ross, *Grand Unified Theories*, (Benjamin-Cummings Pub C, 1984).
- [6] M. Baak, M. Goebel, J. Haller, A. Hoecker, D. Ludwig, K. Moenig, M. Schott, J. Stelzer, “Updated Status of the Global Electroweak Fit and Constraints on New Physics,” arXiv:1107.0975v1, 2011.
- [7] P. Krieger, The Search for Supersymmetry, 2000,
<http://www.physics.utoronto.ca/~krieger/talks/SUSYColloquium.pdf>.
- [8] CDF Collaboration, “Measurement of the Inclusive Jet Cross Section at the Fermilab Tevatron p-pbar Collider Using a Cone-Based Jet Algorithm,” arXiv:0807.2204v4, 2009.
- [9] M. Wilers, Parton Distributions at the Dawn of the LHC, 2010,
http://www.stfc.ac.uk/PPD/resources/PowerPoint/yr10.lecture09_wielers.ppt.
- [10] BBC World News, Guide to the Large Hadron Collider , 2009,
<http://news.bbc.co.uk/2/hi/science/nature/7567926.stm>.
- [11] A. Bettini, *Introduction to Elementary Particle Physics*, (Cambridge University Press, 2008).
- [12] W. R. Leo, *Techniques for Nuclear and Particle Physics Experiments*, (Springer-Verlag Berlin Heidelberg, 1987).
- [13] CMS Collaboration, CMS WorkBook: Root Basics, 2012,
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookBasicROOT>.
- [14] CMS Collaboration, CMS WorkBook: Physics Analysis Oriented Event Display, 2011,
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookFireworks>.

- [15] CMS Collaboration, CMS WorkBook: CMS Computing Model, 2011,
<https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookComputingModel>.
- [16] CMS Collaboration, “MET Performance in the 2011 Data,” CMS DP-2011/010 (2011).
- [17] CMS Collaboration, “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV,” CMS-PAS-PFT-10-002 (2010).
- [18] Tara Scarborough, Kenichi Hatakeyama, Suvadeep Bose, 2012,
<https://indico.cern.ch/contributionDisplay.py?contribId=7&confId=174689>.
- [19] CMS Collaboration, “Missing transverse energy performance of the CMS detector,” JINST **6**, P09001 (2011).
- [20] F. Beaudette, R. Bellan, D. Benedetti et. al, “Fake missing transverse momentum in the RA2 analysis”, CMS Note 2010/383, 2010.
- [21] S. Eno, D. Ferencek, W. Flanagan, L. Gouskos, K. Hatakeyama et. al, “Fake missing transverse momentum in hadronic and leptonic SUSY searches”, CMS Note 11-188, 2012.
- [22] H. Liu , A. Apresyan, P. Janot et al., “Results of a visual scan of high E_T events in 7 TeV pp collision data”, CMS Note 2010/219, 2010.
- [23] CMS Collaboration, “Search for supersymmetry in the multijets and missing transverse momentum final state in proton-proton collisions at $\sqrt{s} = 7$ TeV”, 2012.
- [24] CMS Collaboration, “Search for New Physics with Jets and Missing Transverse Momentum in pp Collisions at $\sqrt{s} = 7$ TeV,” JHEP **2011** (2011).
- [25] CMS Collaboration, MadGraph in CMS, 2009,
<https://twiki.cern.ch/twiki/bin/view/CMS/MadGraphCMSPage>.
- [26] T. Sjostrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 Physics and Manual; v6.420, tune Z2,” JHEP **2006** (2006).
- [27] CMS Collaboration, “Interpretation of Searches for Supersymmetry with Simplified Models,” (2012).
- [28] M. D’Alfonso, S. Mrenna, C. Silkworth, W. Waltenberger, Interpretation of Searches for Supersymmetry with Simplified Models, CMS Note 11-328, CERN, 2012.
- [29] A. L. Read, “Presentation of search results: the CLs technique,” J. Phys. G **28**, 2693 (2002).

- [30] T. Junk, “Confidence level computation for combining searches with small statistics,” Nucl. Instrum. Methods Phys. Res., Sect. A **434**, 435 (1999).
- [31] ATLAS and CMS Collaborations, Procedure for the LHC Higgs boson search combination in Summer 2011, ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, 2011.
- [32] Stefano Forte, CMS WorkBook: CMS Computing Model, 2011, <http://th-www.if.uj.edu.pl/acta/vol41/pdf/v41p2859.pdf>.
- [33] W. Beenakker, R. Hoepker, and M. Spira, “PROSPINO: A Program for the Production of Supersymmetric Particles in Next-to-leading Order QCD,” (1996).
- [34] ATLAS Collaboration, “Hunt for new phenomena using large jet multiplicities and missing transverse momentum with ATLAS in $L = 4.7 \text{ fb}^{-1}$ of $\sqrt{s} = 7$ TeV proton-proton collisions”, Technical Report ATLAS-CONF-2012-037, CERN, Geneva, 2012.