

UNIVERSITY OF CALIFORNIA, SAN DIEGO

**A search for new physics using the M_{T2} variable in all-hadronic final states
produced in proton-proton collisions with a center of mass energy of 13 TeV**

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Physics

by

Mark Derdzinski

Committee in charge:

Professor Frank Würthwein, Chair
Professor Avraham Yagil
Professor Benjamin Grinstein
Professor Farhat Beg
Professor Ian Galton

2018

Copyright
Mark Derdzinski, 2018
All rights reserved.

The dissertation of Mark Derdzinski is approved, and
it is acceptable in quality and form for publication on
microfilm and electronically:

Chair

University of California, San Diego

2018

DEDICATION

FIXME: Dedication here.

EPIGRAPH

For knowledge comes slowly, and when it comes, it is often at great personal expense.

— Paul Auster, *Ghosts*

TABLE OF CONTENTS

Signature Page	iii
Dedication	iv
Epigraph	v
Table of Contents	vi
List of Figures	viii
List of Tables	ix
Acknowledgements	x
Vita	xi
Abstract of the Dissertation	xii
Chapter 1 Introduction	1
1.1 The Standard Model of Particle Physics	1
1.2 Issues with the Standard Model	1
1.3 Beyond Standard Model Physics	1
Chapter 2 The Large Hadron Collider and the CMS Detector	2
2.1 The Large Hadron Collider	2
2.2 The Compact Muon Solenoid	2
2.2.1 Silicon Tracker	4
2.2.2 Electromagnetic Calorimeter	5
2.2.3 Hadronic Calorimeter	9
2.2.4 Muon Detectors	11
2.2.5 Trigger Systems	13
2.3 CMS Physics Objects	15
2.3.1 Particle Flow	15
2.3.2 Isolation	16
2.3.3 Electrons, Muons, and Photons	17
2.3.4 Muons	18
2.3.5 Jets	18
2.3.6 Missing Energy	20

Chapter 3	The M_{T2} Variable Search	22
	3.1 Analysis Strategy	22
	3.2 The M_{T2} Variable	23
	3.3 Event Selection Criteria	25
	3.4 Search Regions	26
	3.5 Control Regions	29
	3.5.1 $\gamma + \text{jet}$ Control Region	29
	3.5.2 Single Lepton Control Region	29
	3.5.3 $Z \rightarrow l^+ l^-$ Control Region	30
	3.5.4 Multijet Control Region	31
Chapter 4	Background Estimates	32
	4.1 Types of Backgrounds	32
	4.2 Multi-jet Estimate	32
	4.3 Lost Lepton Estimate	32
	4.4 Invisible Z Estimate	32
Chapter 5	Results	33
	5.1 Yields and Significances	33
	5.2 Signal Interpretations	33
Chapter 6	Naturalness and Soft Leptons	34
	6.1 Natural Supersymmetry	34
	6.2 Extending the M_{T2} Search	34
	6.3 Searching for Natural Susy with Soft Leptons	34
Bibliography	35

LIST OF FIGURES

Figure 2.1:	Geometry of silicon tracker inner layers in CMS.	5
Figure 2.2:	Feynman diagrams depicting the main processes by which particles shower in the ECAL. The left diagram depicts electron-positron pair production from a photon, and the right diagram depicts bremsstrahlung, where an electron radiates energy away through a photon.	6
Figure 2.3:	Energy resolution σ/E of the ECAL as a function of electron energy measured using a test beam. The energy was measured in a 3×3 crystal array with electrons incident on the center crystal, with electrons falling in a $4 \times 4 \text{mm}^2$ region (lower points) and $20 \times 20 \text{mm}^2$ region (upper points).	7
Figure 2.4:	A cross section of the ECAL geometry, with the dashed lines marking the pseudorapidity values η covered by the various subsystems.	8
Figure 2.5:	The jet transverse energy resolution as a function of the jet transverse energy, in different regions of pseudorapidity. For an explanation of jets, see section 2.3.5	10
Figure 2.6:	Muon momentum resolution as a function of muon momentum using only the inner tracking system, only the muon system, or both combined in the barrel region (left) or endcap region (right).	12
Figure 2.7:	Layout of the CMS muon system for initial run configurations. The DT system is used only in the barrel region and CSCs in the endcap region, while RPCs are deployed in both the barrel and endcap.	12
Figure 2.8:	A graphic depiction of different particles leaving various signatures in the different CMS detector subsystems. Particles may be detected via tracks hits, energy deposits, or a combination of both.	16
Figure 3.1:	Topological regions in N_{jets} and $N_{\text{b-jets}}$ for the $[575, 1000]$ H_{T} region. Within each region, the relative fraction of background events from different SM processes is shown based on simulation.	26
Figure 3.2:	Signal region bins in H_{T} and $E_{\text{T}}^{\text{miss}}$ (left) and $M_{\text{T}2}$ binning within each H_{T} region (right). If simulation predicts less than one background event in the greatest $M_{\text{T}2}$ bin within a region, it is merged with the previous bin.	28

LIST OF TABLES

Table 3.1:	Summary of physics objects and preselection for events.	25
Table 3.2:	M_{T2} binning in the Very Low, Low, and Medium H_T topological regions.	27
Table 3.3:	M_{T2} binning in the High and Extreme H_T topological regions.	27
Table 3.4:	Definition of "super signal regions" used in reinterpretations of the analysis.	29

ACKNOWLEDGEMENTS

FIXME: Acknowledgements

VITA

2011	B. A. in Physics and Mathematics, University of California, Berkeley
2014	M. S. in Physics, University of California, San Diego
2018	Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

Search for new physics with the M_{T2} variable in all-jets final states produced in pp collisions at $\sqrt{s} = 13$ TeV, *The CMS collaboration, Khachatryan, V., Sirunyan, A.M. et al.* J. High Energ. Phys. (2016) 2016: 6. doi:10.1007/JHEP10(2016)006 [arXiv:1603.04053](#) [hep-ex]

Search for new physics in the one soft lepton final state using 2015 data at $\sqrt{s} = 13$ TeV, *CMS Collaboration*, Physics Analysis Summary (2016), CMS-PAS-SUS-16-011

FIXME: Add 2017 MT2 paper citation

ABSTRACT OF THE DISSERTATION

A search for new physics using the M_{T2} variable in all-hadronic final states produced in proton-proton collisions with a center of mass energy of 13 TeV

by

Mark Derdzinski

Doctor of Philosophy in Physics

University of California, San Diego, 2018

Professor Frank Würthwein, Chair

FIXME: Add Abstract.

Chapter 1

Introduction

1.1 The Standard Model of Particle Physics

1.2 Issues with the Standard Model

1.3 Beyond Standard Model Physics

Chapter 2

The Large Hadron Collider and the CMS Detector

2.1 The Large Hadron Collider

FIXME: add paragraphs on LHC

2.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is a general-purpose physics detector at the LHC, situated at one of the five collision points along the main ring. The detector encapsulates the collision point with layers of various subsystems designed to interact with the outgoing particles of the proton-proton collisions, and measure the position and energies of the collision products. Because of the extremely high rate of interactions at the collision

point (on the order of one billion interactions per second), saving data from every bunch crossing would be unsustainable, and so the detector is also equipped with a system of hardware and software implemented "triggers" which identify events of interest for physics analyses to be saved to disk for further analysis.

The physical construction of the detector is motivated by the different interaction of particles with different types of materials, and consists of several subsystems layered as coaxial cylinders around the interaction point. Each subsystem consists of (sometimes different) components covering the fiducial area coaxial with the beamline (the "barrel") and also the ends of the cylinder (the "endcaps"). The innermost subsystem of CMS is the silicon tracker, which consists of many layered silicon pixels designed to pinpoint the locations of charged particles while minimally interacting with the particle's trajectory. The layer beyond the tracker is the electromagnetic calorimeter (ECAL), a grid of lead-tungstate crystals which scintillate to measure the energies of electromagnetic particles. Beyond the ECAL is the hadronic calorimeter (HCAL), a sampling calorimeter designed to measure the energies of hadronic particles (which deposit minimal energy in the ECAL). The final, outer layer is the CMS muon detector, where the muon detection stations are interweaved with the magnetic return yoke that generates the toroidal 3.8T magnetic field inside the detector volume. The total dimensions of the detector are 21.6m long and 14.6m in diameter, weighing over 12,500 tons.

2.2.1 Silicon Tracker

The silicon vertex tracker (SVT) is a series of silicon pixels and strips designed to measure the position of charged particles in the detector, while disturbing their path as little as possible. The position of particles in the interior is of particular importance in event reconstruction; charged particles traveling in a magnetic field will deflect in a curved path with radius proportional to the particles momentum as described in equation 2.1, and so the track reconstruct can be used to not only determine a particle's momentum with high precision, but also the sign of its charge based on the direction of curvature.

$$p = qrB \tag{2.1}$$

As the innermost detector subsystem, the SVT experiences the highest flux of particle radiation. In the barrel region, the tracker layers are oriented in 3 coaxial layers. Closest to the interaction point where particle flux is the greatest, very precise silicon pixels are used, measuring $100 \times 150 \mu\text{m}^2$, whereas in other layers the flux is low enough to use microstrip detectors, measuring $10\text{cm} \times 80 \mu\text{m}$ and $25\text{cm} \times 180 \mu\text{m}$ in the middle and outer layers respectively. In the endcaps, the pixel strips are arranged in a turbine-like pattern in two separate layers on each end. This configuration allows for the precise measurement of particle position for track reconstruction, while minimizing the amount of material which might deflect particles from their original trajectories. A partial geometry of the SVT layout can be seen in figure 2.1.

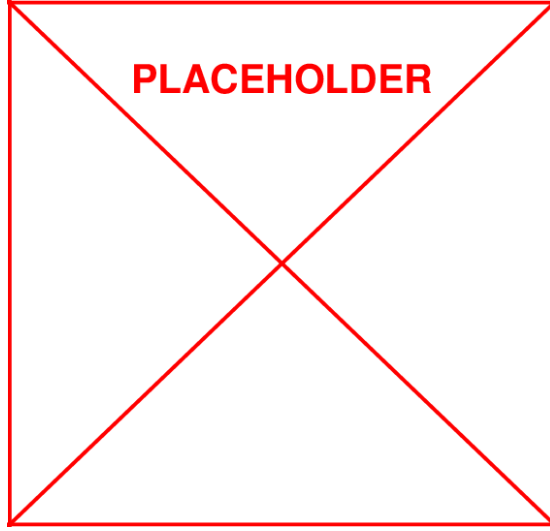


Figure 2.1: Geometry of silicon tracker inner layers in CMS.

2.2.2 Electromagnetic Calorimeter

The ECAL is used to measure the energies of particles which interact electromagnetically, both absorbing the incident particles and scintillating to provide an energy-readout to photodiodes attached to each crystal. Constructed of lead-tungstate (PbWO_4), electromagnetically interacting particles (such as electrons or photons) will interact with the crystal material, losing energy through a cascade of electromagnetic interactions including electron-positron pair production and bremsstrahlung, as pictured in figure 2.2. This phenomenon — also referred to as "showering" — causes the crystals to scintillate proportional to the energy deposited in the crystal, which is then measured by various photodiodes to extract an accurate measurement of the particle energy, now fully absorbed by the calorimeter.

The fundamental principle of the calorimeter measurement relies on the energy loss

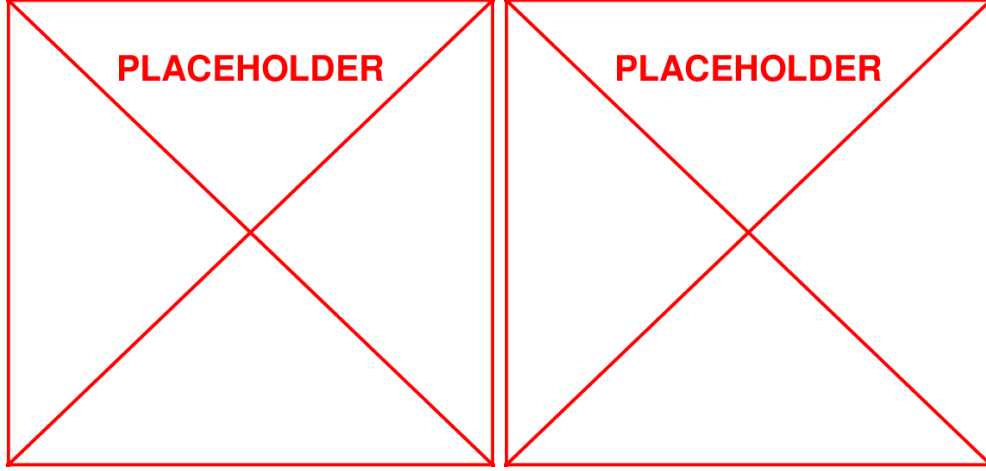


Figure 2.2: Feynman diagrams depicting the main processes by which particles shower in the ECAL. The left diagram depicts electron-positron pair production from a photon, and the right diagram depicts bremsstrahlung, where an electron radiates energy away through a photon.

of particles interacting with matter. In general, the energy of a particle traveling a distance X through some material is given by equation 2.2, where E_0 is the initial energy of the particle and X_0 is the material-dependent radiation length.

$$E(x) = E_0 e^{-\frac{x}{X_0}} \quad (2.2)$$

The design of the calorimeter is motivated by the choice of a scintillating, radiation-hard material with short X_0 such that incident electromagnetic particles deposit all their energy and are stopped by the ECAL. The resolution of the energy measurement is also dependent on the "stochastic term", which parametrizes the uncertainty due to statistical and measurement fluctuations in the calorimeter, and is given by equation 2.3, where S is

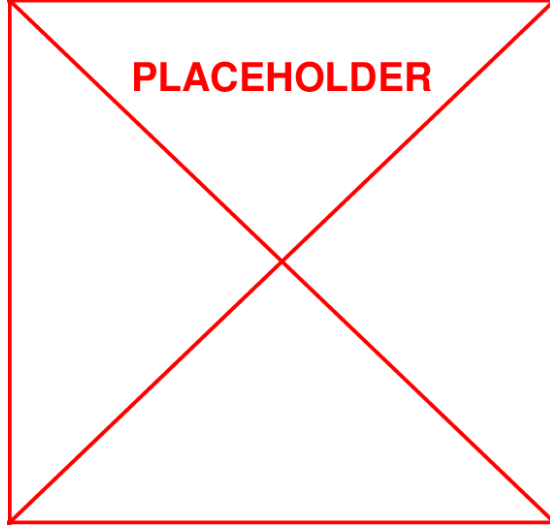


Figure 2.3: Energy resolution σ/E of the ECAL as a function of electron energy measured using a test beam. The energy was measured in a 3×3 crystal array with electrons incident on the center crystal, with electrons falling in a $4 \times 4 \text{mm}^2$ region (lower points) and $20 \times 20 \text{mm}^2$ region (upper points).

the stochastic term, N the noise, and C the constant term.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (2.3)$$

The energy resolution can be measured by a test beam of known energy, as shown in figure 2.3.

The construction of the calorimeter is also divided into two sections by the cylindrical geometry, the ECAL barrel section (EB) and ECAL endcap sections (EE). The EB consists of 61,200 crystals arranged into 36 "supermodules", each spanning half the barrel length, and uses silicon avalanche photodiodes (APDs) as photodetectors. The individual crystals are tilted slightly (3°) in an $\eta - \phi$ grid with respect to the nominal interaction

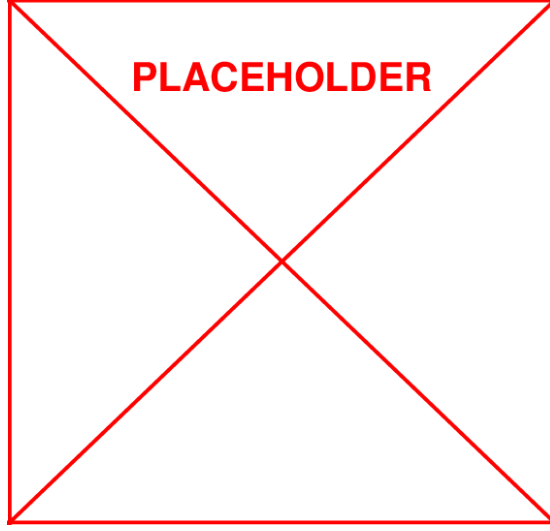


Figure 2.4: A cross section of the ECAL geometry, with the dashed lines marking the pseudorapidity values η covered by the various subsystems.

point, with a front-facing area of $22 \times 22 \text{mm}^2$ and a length of 230mm. The EE instead uses vacuum phototriodes (VPTs) as photodetectors, and consists of approximately 15,000 crystals clustered in 5×5 units, also offset from the interaction point but arranged in an $x - y$ grid, with a cross section of $28.6 \times 28.6 \text{mm}^2$ and a length of 220mm. The EE is also equipped with a "preshower" device placed in front of the crystal calorimeter, consisting of two strips of silicon strip detectors to enhance π^0 rejection. The layout of the ECAL can be seen in figure 2.4. Because of the depth of the ECAL crystals (which are $\sim 25X_0$, and the confining properties of the crystals (which have a Moliere radius of 2.2cm, the radius of a cylinder containing 90% of a shower's energy on average), electrons and photons are typically well reconstructed in CMS, except in the transition region where EB and EE meet.

2.2.3 Hadronic Calorimeter

The CMS HCAL is a sampling calorimeter. Designed with alternating layers of scintillating and absorbing material, incident hadronic particles (such as charged pions, kaons, protons, etc.) interact with the absorber material and consequently shower into electromagnetic particles, whose energy can be read out by photodiodes connected to the scintillating material. Brass is used as the absorber material for both its interaction length and non-magnetic properties, and plastic scintillator tiles connected to embedded wavelength-shifting fibers carry the light to a readout system.

As with the ECAL, the energy loss of hadronic particles in the absorber is characterized by the (hadronic) interaction length and equation 2.2. However, unlike the ECAL, the HCAL contains both hadronic and electromagnetic showers. Electromagnetic particles generated in hadronic showers often fail to escape the absorber layers, and thus some electromagnetic energy is lost in the absorbers. The CMS HCAL is sometimes referred to as a *non-compensating calorimeter* because it is not constructed to actively compensate for the energy lost to these electromagnetic effects and the energy measurements must be corrected offline, known as "jet energy corrections" (JECs). JECs are typically calculated by examining data from collisions producing a boson recoiling against hadronic particles. By accurately measuring the boson energy in the ECAL, the sum of the recoiling hadronic energy in the HCAL is inferred (by momentum conservation) and compared to the detector response. Because the performance of the HCAL can fluctuate with time and run conditions, JECs are regularly recalculated and applied to the raw energy measurements

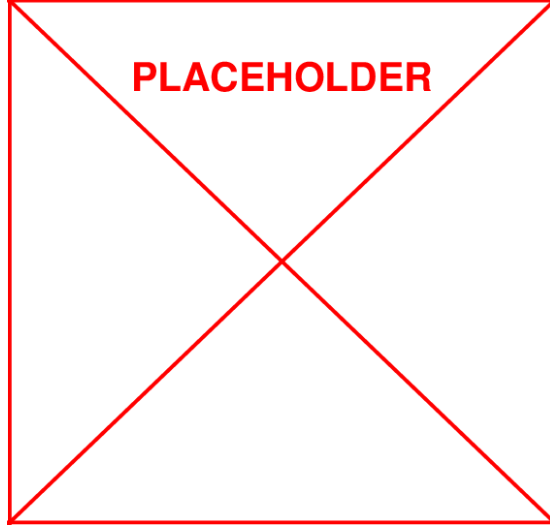


Figure 2.5: The jet transverse energy resolution as a function of the jet transverse energy, in different regions of pseudorapidity. For an explanation of jets, see section 2.3.5

taken by the HCAL to compensate for these effects. Additional information on the corrections contained in JECs is detailed in section 2.3.5. The energy resolution of the HCAL in different regions of pseudorapidity can be seen in figure 2.5.

The geometry of the HCAL can be reduced to four sections. The HCAL barrel (HB) consist of 32 "towers" of alternating absorber/scintillator material spanning the pseudorapidity region $-1.4 < \eta < 1.4$. The HCAL outer (HO) detector lies outside the vacuum tank of the magnetic coil and measures energy from any hadronic particles "leaking" through the HB, covering the pseudorapidity region $-1.26 < \eta < 1.26$. The HCAL endcap (HE) consists of 14 η towers spanning the pseudorapidity range $1.3 < |\eta| < 3.0$. Finally, the HCAL forward (HF) is a different steel/quartz fibre calorimeter spanning the very-forward $3.0 < |\eta| < 5.0$ region. Instead utilizing Cherenkov radiation generated in the quartz fibers, the HF preferentially samples neutral hadronic energy and is ideally designed

for the hadronic-heavy radiation environment in the forward region.

2.2.4 Muon Detectors

The muon detector is the only subsystem which is constructed outside of the toroidal magnetic field. Interleaved with the magnetic return yoke, different muon detectors are used to aid in the identification and reconstruction of muon tracks. Because muons typically penetrate every other layer of the detector and have a large bending radius, additional measurements in the muon system — combined with measurements in the SVT — can lead to vastly improved resolution of muon momentum. The non-uniform magnetic field in the muon detector region (beyond the toroidal regime) causes an s-shaped trajectory and tighter bending radius (than in the SVT) for the muons, which improves the resolution for particles with transverse momentum above $\sim 200\text{GeV}$ as seen in figure 2.6.

Different types of muon detectors are deployed (sometimes in combination) in different sections of the full muon system. In the barrel region ($-1.2 < \eta < 1.2$) where the muon flux, neutral background, and magnetic field are small, drift tube (DT) chambers are used. In the endcaps where the the muon flux, neutral background, and magnetic field are much greater, cathode strip chambers (CSC) are used to increase coverage up to $|\eta| < 2.4$. In addition to these technologies, both the barrel and endcap systems are supplemented with resistive plate chambers (RPC) to provide complementary information to the DT and CSC detectors. The layout of the different muon detector components in the barrel and endcap can be seen in figure 2.7.

The DT detectors are chambers filled with gas surrounding a wire, with a voltage

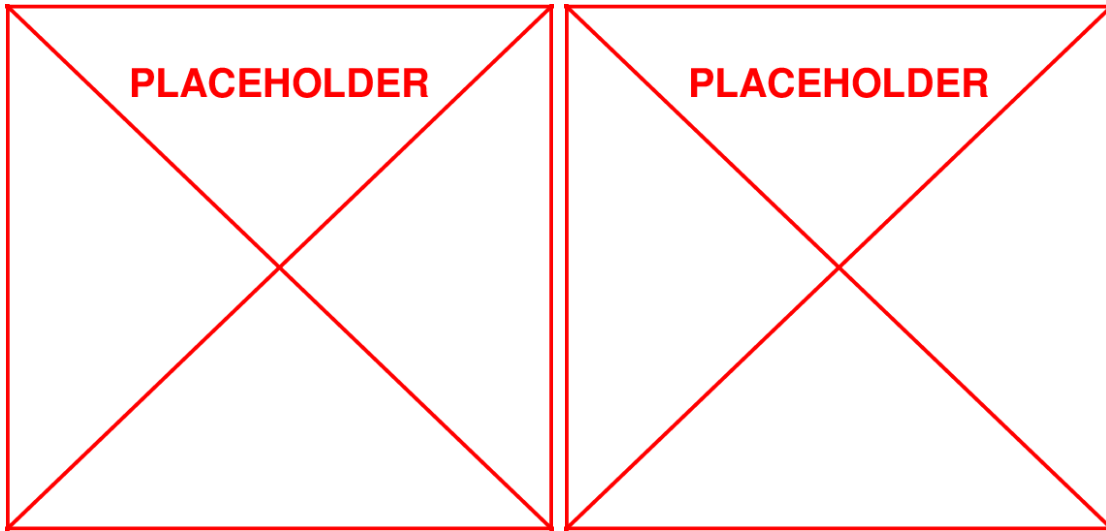


Figure 2.6: Muon momentum resolution as a function of muon momentum using only the inner tracking system, only the muon system, or both combined in the barrel region (left) or endcap region (right).

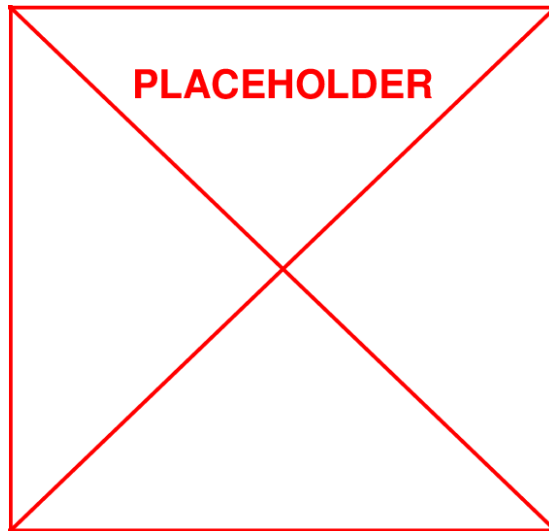


Figure 2.7: Layout of the CMS muon system for initial run configurations. The DT system is used only in the barrel region and CSCs in the endcap region, while RPCs are deployed in both the barrel and endcap.

difference between the wire and outside of the DT. When charged particles pass through the drift chamber, they ionize the gas, and the ionization products will drift across the voltage difference in the tube, resulting in a detectable voltage change in the DT. By measuring both the position along the DT wire where charge is deposited, as well as reconstructing the "drift time" it takes for the ionized particles to reach the wire, DTs provide a 2-dimensional measurement of a particle's position. In CMS, the DT chambers consist of a dozen layers arranged into 3 groups, each with up to 60 DTs. The layers are arranged in such a way that some measure the direction of the muon parallel to the proton beam and others along the perpendicular coordinate, such that the full muon trajectory can be reconstructed by using the DT station information.

FIXME: add paragraph on CSCs

FIXME: add paragraph on RPCs

2.2.5 Trigger Systems

When operating at design luminosity, the LHC can deliver proton bunches to the collision point at a rate of 40MHz (or 25ns between bunches), resulting in an average collision rate on the order of one billion collisions per second. In order to achieve reasonable rates of data collection for offline storage and processing, the detector must suppress the event rate by six orders of magnitude when selecting events of interest to be saved for physics analyses. This is accomplished through a combination of readout electronics and the trigger systems: the Level-1 trigger (L1) processors and online High-Level triggers (HLT).

The L1 trigger system is comprised of specialized hardware processors to rapidly pre-select events of interest based on the calorimeter and muon systems. Based on the beam crossing frequency, the L1 electronics have only a few microseconds to collect readout data from the front end electronics and execute logic to select events of interest, such that the total time allotted for L1 trigger calculations is $< 1\mu s$. During this time the bulk of detector data is held in a buffer while L1 trigger decisions are made based on data with reduced granularity and resolution rapidly collected from the calorimeter and muon systems, where triggers typically check for "trigger primitive" objects (such as photons, muons, electrons, etc.). Trigger primitives must meet certain momentum or energy thresholds, and L1 triggers may also check global data about the event such as the sum of transverse energy or the missing transverse energy (inferred from momentum conservation).

After an L1 trigger tags an event, high-resolution data is read out from buffers for additional data processing before reaching the HLT. Each event of $\sim 1.5MB$ is transferred to front-end readout buffers which then pipes data to a processor containing the HLT code. HLT code is designed to discard events as soon as possible when making trigger decisions and only a subset of objects or partial reconstruction of events may occur before the final trigger decision is made, though HLT algorithms typically approach the quality of final reconstruction. The HLT reduces the L1 output rate to $\mathcal{O}(100\text{Hz})$ for event storage and full reconstruction.

2.3 CMS Physics Objects

When physics events are fully reconstructed, detector data is used to identify *physics objects* representing real particles and event quantities for use in a physics analysis. The physics objects in an event — such as leptons, jets, or missing energy — and their properties are used to select events of interest for physics analyses targeting different final states. The properties of physics objects and global event data are also be used to make analysis level decisions of the quality of different objects. Here we describe some of the physics objects referred to in the M_{T2} analysis and how they are reconstructed, as well as some global event properties and quality variables used as discriminants for physics objects and events.

2.3.1 Particle Flow

Most of the physics objects described in the following sections are reconstructed and identified in CMS using the particle flow (PF) algorithm. The PF algorithm is a holistic, iterative algorithm which uses all the available data in the detector to classify "PF candidate" particles in an event. PF works iteratively by identifying tracks and calorimeter deposits into a PF candidate, removing all energy and hits associated with the candidate and repeating the algorithm until all the detector information has been associated to PF objects. First any muon tracks in the inner tracker associated with muon system hits are associated and removed. remaining tracks are extrapolated into the calorimeters, and any energy deposits on the path are associated with the track and removed from further consideration. Once all the tracks have been associated, the remaining energy clusters can

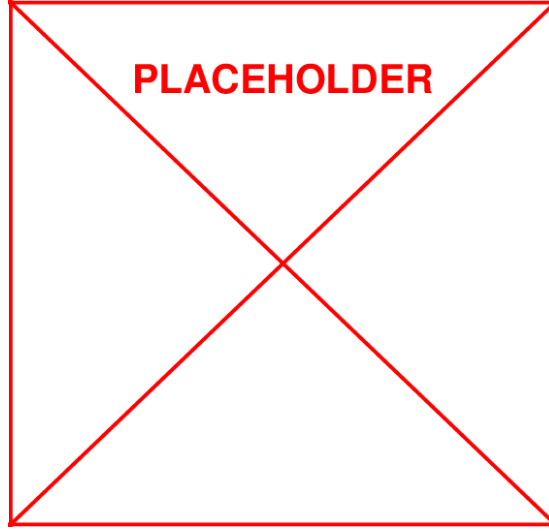


Figure 2.8: A graphic depiction of different particles leaving various signatures in the different CMS detector subsystems. Particles may be detected via tracks hits, energy deposits, or a combination of both.

be identified with photons and neutral hadrons (depending on their presence in the ECAL or HCAL, respectively). An example of the different tracks and energy deposits associated with various particles can be seen in figure 2.8.

2.3.2 Isolation

Isolation is an important distinguishing variable for physics objects measured in the detector. Simply put, isolation measures the total amount of energy in some proximity to a physics object. Because many physics processes results in particle production outside the primary interaction vertex (such as showering, decays, hadronization, pair production, etc.), the isolation of a particle parametrizes the amount of other "activity" surrounding the physics object, and consequently proves a useful discriminant when determining if

particles are produced in the primary interaction, or in subsequent physics processes. Particles which are produced in the primary physics process in a collision are sometimes referred to as *prompt*, and isolation is a powerful tool in identifying prompt leptons in particular.

2.3.3 Electrons, Muons, and Photons

The primary leptons considered in this analysis are electrons and muons (discussed in section 2.3.4), and photons are also used as a cross-check in some control regions for background estimates. Prompt electrons and photons are identified in CMS primarily through the use of the ECAL, where they are stopped and deposit all their energy. Prompt electrons are customarily distinguished by both the presence of a track originating at the primary vertex and a sufficiently low isolation.

Electrons in CMS are identified primarily through the existence of a charged particle track terminating in an energy cluster in the ECAL. As described in section 2.2.2, the ECAL crystals are ~ 25 radiation lengths deep, and will stop and contain nearly all of the energy of an electron (except perhaps miniscule losses to scattering in the tracker). The charged track leading to the energy deposit distinguishes the charged particle from a neutral particle, and the direction of curvature of the path in the tracker can be used to identify the charge of an electron (or positron).

Photons are distinguished from electrons by the lack of a track in the SVT. As a neutral particle, the only sign of a photon will be the energy deposit in the ECAL where the photon is stopped and deposits all its energy. Photons are typically distinguished

from other neutral electromagnetic particles (such as a π^0) by the shape of the shower in the ecal. Whereas photons typically shower by the processes in figure 2.2, other hadronic particles may cascade via different physics processes leading to measurable difference in the shape of the energy cluster.

2.3.4 Muons

Prompt muons produced in collisions at the LHC are generally minimum ionizing particles, and will penetrate the bulk of the detector. Without any matching calorimeter deposits, they are reconstructed using hits in the muon system matched to inner tracker hits. The muon system flags muon candidates, which are then matched to inner tracker hits for the best fit of the muon track. Given the track hits, the transverse momentum and energy of the muon can be calculated by a fit to the track parameters. As with electrons, muons are customarily distinguished by both the presence of a track originating at the primary vertex and a sufficiently low isolation.

2.3.5 Jets

Many particles produced at the LHC are created through the strong interaction, which may produce particles carrying color charge in the final state (i.e. quarks or gluons). Colored particles cannot exist individually due to the phenomenon known as *confinement*, and so via the process of *hadronization* (also referred to as *fragmentation*) they will proliferate in a series of interactions producing additional particle-antiparticle pairs, cascading in a parton shower to form hadronic bound states. When a "bare" quark or gluon is pro-

duced in the primary interaction, the ensuing hadronization results in a stream of tightly collimated hadrons aligned with the trajectory of the original bare particle, referred to as a *jet*.

Jets present themselves in CMS as collimated energy deposits in both the ECAL and HCAL, as well as tracks originating at the primary vertex. Between the two calorimeters, all the electromagnetic and hadronic components of the jet are stopped and measured by the calorimeters, and any leptons (including muons detected in the muon system) which originate in the cone of the jet are assigned to the total jet energy. The jets are reconstructed using the "anti- k_T " algorithm, which clusters jet activity in cones of a fixed radius. The jet is treated as a single physics object with a total energy equal to the sum of its electromagnetic, hadronic, and leptonic components, which by momentum conservation dictates the energy of the prompt quark or gluon that was produced in the primary interaction.

Furthermore, the substructure of jets and their content can be analyzed to determine the flavor of the parent parton, known as "tagging". Jets are typically tagged to distinguish between those originating from heavy-flavor quarks (bottom or charmed) and light-flavor quarks. Top quarks produced in primary interactions are unique because of their short lifetime. A top quark will decay before hadronization occurs, instead producing a W boson and down-type (down, bottom, or strange) quark which will subsequently hadronize. Searches involving top quarks in the final state typically employ "top-taggers" to search for these signatures of the top quark.

In practice, jets are complicated objects consisting of many constituent particles,

and must be clustered and calibrated so that the jet energy closely matches that of the parent parton which produced the jet. JECs are applied to raw jet energies to compensate for different experimental deviations from the parent parton energy. In particular, JECs include corrections to compensate for pile-up energy, detector effects (as a function of η), energy scale as a function of jet p_T , and residual corrections to account for differences between data and simulation. The JECs are calculated by collecting data events with a Z boson (decaying to electrons or muons) or photon recoiling against a jet. By precisely measuring the bosons energy with the ECAL, scale factors are derived for jets as a function of pseudorapidity η and momentum p_T .

2.3.6 Missing Energy

Missing energy refers to the sum of all energy which has escaped the detector, and is inferred from a momentum-imbalance in the physics objects measured by the detector. Because the momentum of the incident protons is zero perpendicular to the direction of the beam, the momentum of physics objects produced in any collision must sum to zero in the direction transverse to the beamline. The missing transverse energy (E_T^{miss}), is calculated by taking the negative vector sum of all PF candidates as in equation 2.4. E_T^{miss} is of particular importance to analyses targeting BSM physics, as BSM signatures characteristically contain invisible particles in the final state which escape detection.

$$\vec{E}_T^{\text{miss}} = - \sum_i \vec{p}_T^i \quad (2.4)$$

The missing energy in an event is inferred from all other measured quantities in an

event, and there are many sources of E_T^{miss} which are unphysical in nature but rather dependent on experimental effects. Particles from the primary interaction with a sufficiently large pseudorapidity may escape the fiducial region of the detector subsystems, and resolution effects or intrinsic noise in the detector may lead to fluctuations in measured energies. These experimental effects must be suppressed or distinguished from "real" E_T^{miss} due to physics processes creating particles which escape the detector (e.g. neutrinos). Analyses sensitive to final states with E_T^{miss} often employ robust data-driven methods to predict or suppress backgrounds which might generate experimental sources of E_T^{miss} or physics processes which can contribute real E_T^{miss} (such as $Z \rightarrow \nu\nu$).

Chapter 3

The M_{T2} Variable Search

3.1 Analysis Strategy

Searches for new physics targeting all-hadronic final states present unique challenges and opportunities at the LHC. While such searches typically implement stringent vetoes on lepton candidates and thus bypass the need to correctly identify "real" leptons, the high rate of QCD processes in proton-proton collisions generates large amounts of SM events with all-hadronic final states. Designing a search targeting signatures with all-hadronic final states requires a mechanism to distinguish and suppress the selection of multi-jet QCD events from new physics signatures, as well as robust background estimation methods to predict the yield of standard model events which may generate missing energy (such as $Z \rightarrow \nu\nu$).

The M_{T2} analysis harnesses the discriminating power of the M_{T2} variable to distinguish standard model events from possible signatures including new physics. By first

requiring a significant amount of missing energy in the event, multi-jet QCD processes are greatly suppressed. Additional requirements on the topology of the event implemented using M_{T2} further suppress QCD-like processes and favor events with real missing energy anti-aligned with the hadronic energy deposits in the detector. After estimating the minimal QCD contribution remaining by extrapolating from a region orthogonal to the signal selection, the only remaining backgrounds are leptonic events where the lepton failed reconstruction or identification (or "lost-lepton" events), and SM events creating real E_T^{miss} , in the form of neutrinos from a decaying Z boson recoiling against jets (or "invisible Z" events).

3.2 The M_{T2} Variable

M_{T2} is a particularly useful kinematic mass variable for final states where two particles decay (possibly in a chain) to a final state containing an invisible particle X of mass m_X . The typical transverse mass M_T , is defined in equation 3.1 for particles $i = 1, 2$, where the mass $m^{\text{vis}(i)}$, transverse momentum $\vec{p}_t^{\text{vis}(i)}$, and transverse energy $E_T^{\text{vis}(i)}$ characterize the visible kinematics of the decay chain, and $\vec{p}_t^{\text{X}(i)}$ and $E_T^{\text{X}(i)}$ characterize the unknown kinematics of the invisible particle X.

$$(M_T^{(i)})^2 = (m^{\text{vis}(i)})^2 + m_X^2 + 2 \left(E_T^{\text{vis}(i)} \cdot E_T^{\text{X}(i)} - \vec{p}_t^{\text{vis}(i)} \cdot \vec{p}_t^{\text{X}(i)} \right) \quad (3.1)$$

In principle, if the correct values of m_X and $\vec{p}_t^{\text{X}(i)}$ were accessible, then the transverse mass would have a kinematic endpoint and not exceed the mass of the parent particles. However, the individual momenta of the invisible particles in the two decay chains cannot

be measured; the only quantity experimentally accessible is the total missing momentum \vec{p}_T^{miss} . With this in mind, the generalized transverse mass variable M_{T2} is defined in equation 3.2, where the unknown mass m_X is a free parameter and a minimization is performed over the sum of invisible momenta $\vec{p}_t^{\text{X}(i)}$ that satisfy the measured \vec{p}_T^{miss} constraint.

$$M_{\text{T2}}(m_X) = \min_{\vec{p}_t^{\text{X}(1)} + \vec{p}_t^{\text{X}(2)} = \vec{p}_T^{\text{miss}}} \left[\max \left(M_T^{(1)}, M_T^{(2)} \right) \right] \quad (3.2)$$

Because this analysis selects final states with (often several) jets in the final state, the calculation of M_{T2} first requires grouping the hadronic jet activity into two large *pseudojets* to act as the visible components in the M_{T2} equation. The jet activity in each event is divided into two hemispheres, and the jets in each hemisphere are summed together to create the pseudojets. The hemisphere algorithm (as defined in ??) proceeds as follows:

- The direction of the two jets with largest invariant mass is chosen as the initial seed for the two axes.
- Jets are associated to one of the two axes according to the minimal Lund distance, such that jet k is associated to hemisphere i instead of j if the condition in equation 3.3 is true.
- After each jet is associated to one of the two axes, the axes are recalculated by summing the momenta of all jets associated to an axis.
- The association algorithm iterates using the new axes, and continues until no jets are associated to a different axis after an iteration.

FIXME: Table of preselection cuts from MT2 paper

Table 3.1: Summary of physics objects and preselection for events.

$$(E_i - p_i \cos \theta_{ik}) \frac{E_i}{(E_i + E_k)^2} \leq (E_j - p_j \cos \theta_{jk}) \frac{E_j}{(E_j + E_k)^2} \quad (3.3)$$

QCD multijet events (when clustered using this pseudojet algorithm) may yield high M_{T2} events if the pseudojets have high jet masses, thus the visible masses $m^{\text{vis}(i)}$ are set to zero to suppress such SM events. Since the kinetic components of M_{T2} will be large for signal events, this suppression does not significantly impact sensitivity to many BSM signatures, thus M_{T2} is calculated in this analysis using only E_T^{miss} and the two pseudojets.

3.3 Event Selection Criteria

The general strategy for the event selection is to first apply baseline cuts on motivated by hardware and software-level triggers (discussed in section 2.2.5) and reducing the QCD multi-jet background to negligible levels. Events are further categorized using the scalar sum of the transverse momenta p_T of all selected jets (H_T), the total number of jets in the event (N_{jets}), the total number of b-tagged jets in the event ($N_{\text{b-jets}}$), and M_{T2} . A summary of the event preselections can be found in table 3.1.

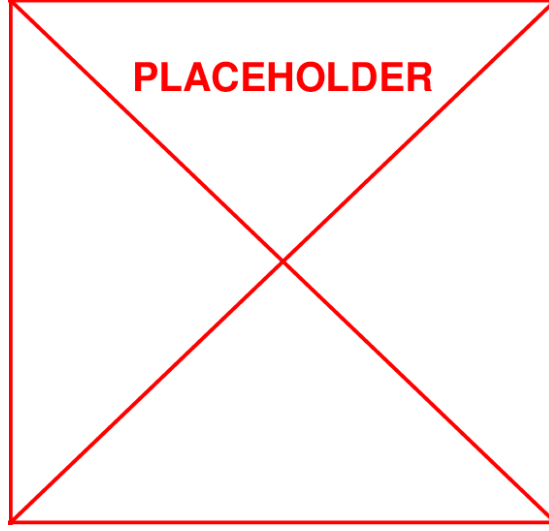


Figure 3.1: Topological regions in N_{jets} and $N_{\text{b-jets}}$ for the $[575, 1000]$ H_{T} region. Within each region, the relative fraction of background events from different SM processes is shown based on simulation.

3.4 Search Regions

The search regions are defined by categorizing events in bins of H_{T} , N_{jets} , $N_{\text{b-jets}}$, and M_{T2} (in addition to the baseline selection described in section 3.3). First events are categorized into "topological regions" according to H_{T} , N_{jets} , and $N_{\text{b-jets}}$:

- H_{T} (GeV): $[250, 450]$ (Very Low), $[450, 575]$ (Low), $[575, 1000]$ (Medium), $[1000, 1500]$ (High), $[1500, \infty]$ (Extreme)
- N_{jets} & $N_{\text{b-jets}}$: 2-3j 0b, 2-3j 1b, 2-3j 2b, 4-6j 0b, 4-6j 1b, 4-6j 2b, $\geq 7\text{j}$ 0b, $\geq 7\text{j}$ 1b, $\geq 7\text{j}$ 2b, 2-6j $\geq 3\text{b}$, and $\geq 7\text{j}$ $\geq 3\text{b}$ (except in the region with $250 < H_{\text{T}} < 450$ GeV, where bins $\geq 7\text{j}$ are merged with 4-6j bins due to lack of events).

The different topological regions for one H_{T} region and their background composition are

FIXME: Table of signal region binning VL, L, M

Table 3.2: M_{T2} binning in the Very Low, Low, and Medium H_T topological regions.

FIXME: Table of signal region binning H, VH

Table 3.3: M_{T2} binning in the High and Extreme H_T topological regions.

depicted in figure 3.1. Each topological region is further divided into bins based on M_{T2} . The M_{T2} binning is constructed such that the low edge of the first bin is 400 GeV in regions with $H_T > 1500$ GeV and 200 GeV everywhere else, and the low edge of the final bin is constructed to contain approximately one background event based on simulation and not exceeding the maximum H_T value in that topological region (since an upper limit on H_T places an upper limit on M_{T2}). The detailed M_{T2} binning is as follows:

- Very Low H_T : [200,300], [300,400], [400, ∞]
- Low H_T : [200,300], [300,400], [400,500], [500, ∞]
- Medium H_T : [200,300], [300,400], [400,600], [600,800], [800, ∞]
- High H_T : [200,400], [400,600], [600,800], [800, 1000], [1000, 1200], [1200, ∞]
- Extreme H_T : [400,600], [600,800], [800,1000], [1000,1400], [1400, ∞]

The various H_T bins and associated M_{T2} binning can be seen in figure 3.2, and the full breakdown of signal regions (including M_{T2} binning) is listed in tables 3.2 and 3.3. In addition to multi-jet search regions, this analysis also considers monojet events. Because

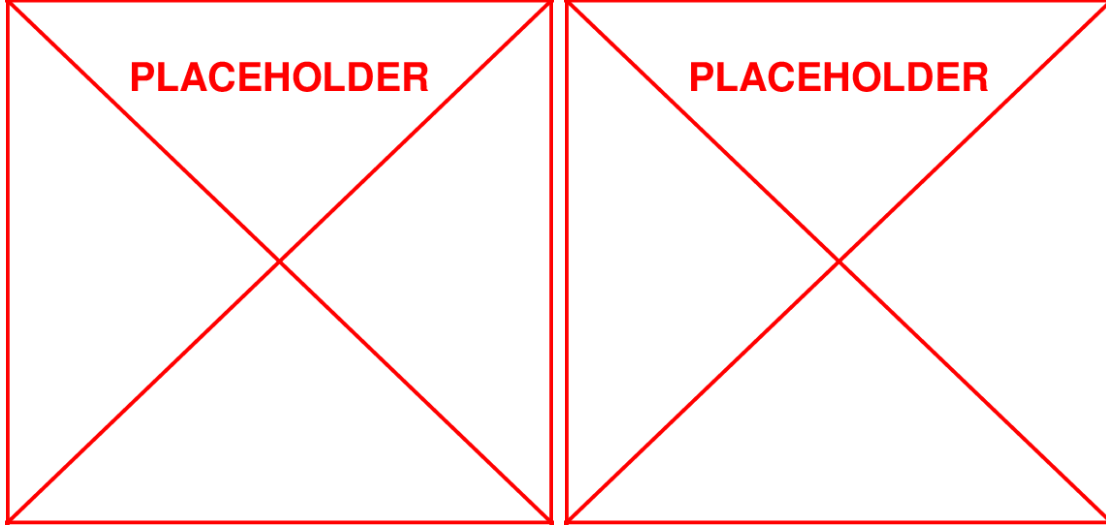


Figure 3.2: Signal region bins in H_T and E_T^{miss} (left) and M_{T2} binning within each H_T region (right). If simulation predicts less than one background event in the greatest M_{T2} bin within a region, it is merged with the previous bin.

there is only a single jet (and M_{T2} is ill-defined without multiple jets), binning in these regions is defined using $N_{\text{b-jets}}$ and H_T as follows:

- $N_{\text{b-jets}}$: 0b, $\geq 1\text{b}$
- H_T : [250,350], [350,450], [450,575], [575,700], [700,1000], [1000,1200], [1200, ∞]

As with the multi-jet regions, monojet H_T bins with less than one simulated background event in the final bin are merged with the penultimate bin.

In addition to these signal regions used to interpret results in the context of various BSM physics models, the analysis also provides results in "super signal regions" (SSRs) as defined in table 3.4. These regions provide a simpler set of selection than the nominal signal regions so that phenomenologists may easily reinterpret results in the context of

FIXME: Table of super signal regions

Table 3.4: Definition of "super signal regions" used in reinterpretations of the analysis.

different signal models (as in **FIXME: reference paper using reinterpretations**).

Results obtained using the SSRs are not as sensitive as the nominal binning — finely binned regions have a higher signal-to-background ratio and the global background fit reduces the background uncertainties – but are much easier to use for reinterpretation than the many correlated bins of the nominal analysis.

3.5 Control Regions

In order to anchor the data-driven background estimates used in this analysis, *control regions* (CR) orthogonal to the signal region selection are defined for various processes. In particular, there are control regions corresponding to enriched samples of $\gamma + \text{jet}$ events, single lepton events, $Z \rightarrow l^+l^-$ events, and QCD multijet events.

3.5.1 $\gamma + \text{jet}$ Control Region

FIXME: photon plus jets CR selection

3.5.2 Single Lepton Control Region

The single lepton CR is constructed to select a sample enriched with single lepton events, the most dominant contributions being from $t\bar{t}$ and $W + \text{jets}$ production. The same

baseline selections described in section 3.3 are applied with the exception of the following:

- In lieu of the lepton veto, exactly one lepton candidate passing the reco or PF lepton selections is required. In order to avoid double counting (for leptons which are reconstructed both as a reco lepton and PF candidate), PF leptons within $\Delta R < 0.1$ of a reco lepton are not considered.
- The transverse mass M_{T2} between the lepton and E_T^{miss} must be less than 100 GeV to reduce possible signal contamination.

Since non-isolated leptons in the fiducial region of the detector are usually successfully reconstructed, the closest jet within $\Delta R < 0.4$ of the lepton is removed and the lepton instead counted as a visible object for the purposes of computing H_T , H_T^{miss} , $\Delta\phi(j_{1234}, E_T^{\text{miss}})$, $|\vec{H}_T^{\text{miss}} - \vec{E}_T^{\text{miss}}|/E_T^{\text{miss}}$, and M_{T2} (as well as the hemispheres used to calculate M_{T2}). Events are further subdivided into the topological regions described in section 3.4 using the modified H_T and N_{jets} and $N_{\text{b-jets}}$, but not in M_{T2} to increase the statistical power of the CR. The signal regions with $\geq 7j, \geq 1b$ are all predicted using CRs with identical H_T bins but $\geq 7j, 1-2b$ to increase the statistical power of those CRs (and to avoid signal contamination in regions with $\geq 7j, \geq 3b$). The monojet CR is binned identically to the signal region.

3.5.3 $Z \rightarrow l^+l^-$ Control Region

FIXME: $Z \rightarrow l^+l^-$ CR selection

3.5.4 Multijet Control Region

FIXME: qcd CR selection

Chapter 4

Background Estimates

4.1 Types of Backgrounds

4.2 Multi-jet Estimate

4.3 Lost Lepton Estimate

4.4 Invisible Z Estimate

Chapter 5

Results

5.1 Yields and Significances

5.2 Signal Interpretations

Chapter 6

Naturalness and Soft Leptons

6.1 Natural Supersymmetry

6.2 Extending the M_{T2} Search

6.3 Searching for Natural Susy with Soft Leptons

Bibliography