UNIVERSITY OF CALIFORNIA
RIVERSIDE

Search for Gluino-Mediated Supersymmetry in Events With Bottom-Quark Jets and Missing Transverse Energy With the Compact Muon Solenoid Detector at the Large Hadron Collider With Proton-Proton Collisions at 8 TeV

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Physics

by

Harold Nguyen

August 2013

Dissertation Committee:
Profesor Bill Gary, Chairperson
Professor Stephen Wimpenny
Distinguished Professor Gail Hanson

The Dissertation of Harold Nguyen is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

There are many people I wish to thank, who helped make the work described in this thesis possible and my graduate school experience more enjoyable.

I start by thanking my main advisor, Bill Gary, who is a great leader. His clarity of thought and honesty are unmatched by anyone I know. He taught me, through example, how to carry myself professionally and keep an optimistic attitude. His explanations are always very clear and precise, and I am very lucky to be his student. I also thank him for giving me the opportunity to study and live in Europe. I should also not forget to thank him for teaching me how to drive a manual transmission, which is just one of his many kind acts.

I also thank Owen Long, who was my co-advisor for the analysis in this thesis. He is a master at his craft. I thank him for his patience in explaining difficult topics to me, helping me with programming code, and being supportive in general throughout my graduate school career.

I thank Steve Wimpenny for his mentorship throughout my graduate career, and for aiding me while I was at CERN. I also thank Bob Clare, John Ellison, and Gail Hanson for their support, and for answering questions I had in particle physics. I thank Jared Sturdy and Sudan Paramesvaran for helping me with the photon studies described in this thesis. I thank Dick Kellogg for teaching me about HCAL front-end electronics.

I would like to thank Marjorie Shapiro from my undergraduate years for telling me about Bill Gary and Owen Long. She is the reason I joined their group. I thank Peter Lost-

cutoff for helping me choose Experimental Particle Physics as the field to study in graduate school. I also thank George Parks and Matt Fillingim for introducing me to physics research, and for being great mentors.

I thank my fellow graduate students from UCR and elsewhere, who made the graduate school experience more fun: Gene, Heman, Yeming, Hang, Feihu, Gil, Hulya, Shih-Chuan, John B, Arun, Pawandeep, Rachel, Mark, and Jeremy. I especially thank David Drummond for the many Chipotle lunches and great friendship, and Haryo Sumowidagdo for his kindness while I was at CERN.

I am tremendously indebted to the RA2b group. I thank Jim Smith, Bill Ford, Keith Ulmer, Ale Gaz, Jeff Richman, Tom Danielson, Kristen Flowers, Anders Ryd, Josh Thompson, Luke Winstrom, Don Teo and Ben Kreis. This group works very well together, and I was lucky to have been a part of it. I would particularly like to thank Keith Ulmer for working with me to measure the trigger performance described in this thesis, and for his general kindness and helpfulness in any question I had.

I thank Teruki Kamon and Anwar Bhatti for their support during the beginning of my graduate career, while I was a member of the LPC JetMET group working on the "SUSY Higgs to bbbar" studies.

I thank Kim for providing endless encouragement and moral support. Her presence brings more joy to my life than anything I can imagine.

Finally, I thank my parents for their unconditional love and support throughout my graduate school experience, and whole life. They have always been there for me.

Dedicated to my parents, who always allowed me to follow my passions.

ABSTRACT OF THE DISSERTATION

Search for Gluino-Mediated Supersymmetry in Events With Bottom-Quark Jets and Missing Transverse Energy With the Compact Muon Solenoid Detector at the Large Hadron Collider With Proton-Proton Collisions at 8 TeV

by

Harold Nguyen

Doctor of Philosophy , Graduate Program in Physics
University of California, Riverside, August 2013
Profesor Bill Gary, Chairperson

A search is presented for physics beyond the standard model based on events with significant missing transverse energy, at least three jets, and at least one identified bottom-quark jet. The study is based on a sample of 19 fb^{-1} collected at 8 TeV with the CMS detector at the Large Hadron Collider in 2012. The background from standard model processes is evaluated using data control samples, and a global likelihood fit is performed. The data are found to be consistent with standard model processes, and the results are interpreted in the context of simplified models (SMS). Upper limits on the production cross sections of the T1bbbb and T1tttt SMS new physics scenarios are determined. Gluino masses up to 1170 GeV are excluded for the T1bbbb scenario and up to 1020 GeV for the T1tttt scenario, at 95% confidence level.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

An elementary particle is a basic unit of matter that, in our current description of particle physics, is treated as being point-like, with no substructure. The standard model (SM) is a quantum field theory that attempts to describe the set of all known elementary particles and their interactions. A more detailed description of the SM is found elsewhere (see Ref. [45]), but a brief summary is given here.

Elementary particles either have spin $\frac{1}{2}$, which makes them fermions, or integral spin, which makes them bosons. Elementary particles that are bosons either have spin 1 (the gauge bosons) or spin 0 (the Higgs boson). Fermions are further split into two classes: quarks and leptons. Quarks carry color charge and thus interact via the strong force. In the SM, color charge is represented through $SU(3)$, which means that the quantum numbers of the quarks do not transform under this gauge group. Quarks also carry electric charge and weak isospin and so interact via the electromagnetic and weak forces, respectively. In

the SM, weak hypercharge (related to the electric charge) is represented through the $U(1)$ gauge group, while weak isospin is represented through the $SU(2)$ gauge group.

Quarks are further classified into three generations of pairs: (up [u], down [d]), (charm [c], strange [s]), and (top [t], bottom [b]). The first (second) element in each pair carries an electric charge of $+\frac{2}{3}e$ ($-\frac{1}{3}e$), where e is the electric charge of a proton. Each generation differs through their quantum numbers, and increasing generations have increasing mass. The quantum numbers of quarks include baryon number and weak isospin. The first generation, which are the up and down quarks, are known to be stable. The second and third generations are unstable and decay to the first generation. Protons, which are the particles being collided in the experiment described in this thesis, are made from a triplet combination of quarks (uud), known as the valence quarks. In addition, protons contain quarks of all flavors and gluons, which are created through virtual processes.

Leptons carry weak isospin and are similarly classified into three generations: electron [e], muon [μ], and tau [τ]. They are respectively paired with their accompanying almost massless neutrinos: electron neutrino [ν_e], muon neutrino [ν_μ], and tau neutrino [ν_τ]. While the e, μ, and τ carry an electric charge of -1 and thus interact electromagnetically, the ν_e, ν_μ, and ν_τ carry no electric charge and only interact via the weak interaction.

Elementary particles with spin 1 (gauge bosons) are sometimes referred to as "force carriers" since they mediate the interactions between elementary particles. Each gauge boson is represented by generators of the group whose interaction they are assigned to. For instance, the $SU(3)$ group has eight generators, so the strong interaction is mediated by

eight neutral, color-carrying gauge bosons called gluons. The group $U(1)$ has one generator corresponding to one boson called a B^0 and the $SU(2)$ group has three generators, corresponding to the $W^{\pm}$ and W^0 gauge bosons. The B^0 and W^0 have the same quantum numbers, and mix to form the mass eigenstates of the photon γ and Z^0 boson. Only left-handed particles participate in the $SU(2)$ weak isospin.

The SM is based on a direct group product, which is $SU(3) \times SU(2) \times U(1)$. This implies therefore that all the terms in the Lagrangian are invariant under this symmetry and therefore all gauge bosons should be massless, like the photon. However, the $W^{\pm}$ and Z^0 gauge bosons are observed to be massive. To give these gauge bosons mass, the SM introduces the so-called "Higgs mechanism," which spontaneously breaks the symmetry to the subgroup $U(1)_{em} \times SU(3)$ and introduces a spin-zero elementary particle called a Higgs boson.

The success of the SM is expressed in its ability to account for essentially all particle-physics data including the spectrum of observed particles. Although the Higgs boson remains the last remaining particle to be found, a Higgs-like boson was discovered by two experiments about one year prior to the time of this writing [38, 35]. The two experiments are A Toroidal LHC Apparatus (ATLAS) and the Compact Muon Solenoid (CMS), which are general-purpose particle physics detectors located at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) in Geneva, Switzerland. This thesis is based on analysis of data collected with the CMS detector.

Although the SM has proved successful up to the highest-probed energies (see Ref. [11]), the SM is generally viewed to be incomplete. The SM does not offer an explanation of so-called "dark matter" [2]. Dark matter is hypothesized to account for a large fraction (roughly 84.5%) of total matter in the universe. Its existence was inferred by investigating massive structures in the universe and observing how the expected visible matter's gravitational effects differed from the actual measured gravitational effects. For instance, if only the visible mass were present, the rotational velocity of structures far from the galaxy's center should be slower than what is observed. The relatively high rotational velocity of structures far from a galaxy's center can be accounted for by dark matter. Furthermore, gravitational lensing, which is a phenomena where light bends around a massive object due to gravity, is observed to be stronger than what is expected for visible matter. The presence of dark matter would offer a simple explanation. Dark matter has not been seen directly and is thus proposed to not interact electromagnetically, hence its name, and it has not been observed to interact with the strong force. Many researchers believe that dark matter is composed of weakly interacting massive particles (WIMPs), which interact only through the weak force and the gravitational force. If dark matter can be produced in experiments like CMS, they would then leave a signature in the detector similar to that of a neutrino, and provide large values of so-called missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Section 3.4.6).

The SM does not provide an elegant unification of the strong, weak, and electromagnetic forces. Though unification is not guaranteed, a trend in physics has been to unify these forces in a so-called grand-unified theory (GUT) [43, 13]. Furthermore, the SM does

not address gravity or the so-called "hierarchy problem," which states that the corrections to the Higgs mass require fine tuning and thus seem too highly contrived to be "natural." The hierarchy problem is also referred to as the *naturalness* or *fine-tuning* problem.

Many theories or extensions to the SM have been proposed to address these issues. One very popular theory is Supersymmetry (SUSY) (see Refs. [58, 44, 54, 67, 68, 69, 41, 55]), which predicts the unification of the electromagnetism, weak, and strong couplings at the GUT scale. Exact invariance under SUSY implies that each particle has a super partner, (sparticle) with the same quantum number and mass, and with spin different by one-half. For example, gluinos and squarks, which are the superpartners of gluon and quarks, respectively, have spin $\frac{1}{2}$ and 0, again respectively. If superparticles had the same masses as their corresponding particles, they already would have been discovered. Since they have not been discovered, their masses (if SUSY is correct) must be larger than those of the particles, and indeed beyond what has been experimentally accessible, so the SUSY symmetry must be broken. It is assumed that this SUSY breaking is "soft," which means that although the masses of a particle and its corresponding antiparticle are not the same, their coupling strengths and other properties are identical.

In a generic SUSY model, a certain quantum number of quarks and leptons, namely baryon number and lepton number, respectively, need not be conserved in some interactions. Since conservation of these quantities is observed in nature, R-parity symmetry is

introduced, which disallows these interactions. R-parity is given by:

$$R = (-1)^{3(B-L)+2s} \tag{1.1}$$

where s is the spin, B is the baryon number, and L is the lepton number. SM have particles with $R = +1$, where SUSY particles have $R = -1$.

SUSY offers a "dark matter" candidate, since R-parity-conserving SUSY models [40] cascading to a "lightest supersymmetric particle" (LSP), which is stable and weakly interacting, are typically produced with SM objects such as leptons and hadronic jets (see Section 3.4.3). Furthermore, R-parity ensures that each sparticle decays to an odd number of LSPs. LSPs can escape the detector in an experiment without interacting, but can be observed indirectly through momentum imbalance.

SUSY solves the "hierarchy problem" because the fermion and boson loop corrections to particle masses cancel one another almost completely, thus stabilizing the calculation of the Higgs boson mass. For this stabilization to be "natural" [39, 4, 57], top and bottom squarks are required to be relatively light. This enhances their production either through direct pair production or production mediated by gluinos. The latter is more likely if the gluino production cross section is large. Top and bottom squarks subsequently decay to bottom quarks, so natural SUSY models involve an abundance of bottom-quark jets (see Section 3.4.3).

This thesis is based on data collected by the CMS detector in 2012, corresponding to 8 TeV proton-proton collisions with an integrated luminosity of $19\,\text{fb}^{-1}$. The analysis discussed in this thesis is based on a search for SUSY with missing transverse energy and bottom-quark jets in the final state. It was performed in a collaboration of four research universities. The universities were UC Riverside, UC Santa Barbara, U. Colorado at Boulder, and Cornell U. There are about three to four people per university who participated in the analysis described in this thesis, which consisted of a mixture of grad students, post-docs and professors. This group formed the so-called Reference Analysis 2 b (RA2b) group, where the "b" stands for b-quark jets.

This thesis follows the analysis in Ref. [31], which was performed by the RA2b group using the 2012 dataset of the CMS detector. For this analysis, I contributed to producing common ntuples (Section 3.6), selected the triggers (Section 3.7.1) for the analysis and determined their performance using data-based techniques, and investigated use of a photon control sample to estimate the $\text{Z} \rightarrow \nu\bar{\nu}$ background (Section 4.1.1.) The important work of measuring the trigger performance played a critical role in the analysis. I also spent one year at CERN from March 2009 until April 2010. During this time, I helped monitor, maintain, and improve the HCAL subdetector (Section 2.2.3) on the CMS experiment (Section 2.2), and participated in the so-called "HCAL Prompt Task Force" (Section 2.2.3). After returning from CERN, I was involved in validating jet algorithms (Section 3.4.3). The experiment and analysis methods performed in Ref. [31] will be summarized in this thesis, but the emphasis will be placed on my personal contributions.

Chapter 2

Experiment

This chapter discusses the LHC and the CMS experiment. The LHC provided 8 TeV proton-proton collisions in 2012, which were recorded by the CMS experiment. The CMS experiment consists of different subdetector systems, each playing a crucial and complimentary role to the whole experiment. The different subdetectors are described below.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC), located at the CERN laboratory near Geneva, Switzerland, is the largest particle accelerator in the world. It is situated in a 27 km long circular tunnel, which was originally built to house the LEP accelerator (LEP, the large electron-positron project, which operated between 1989 and 2000, is the highest energy e^+e^- accelerator ever built.). This tunnel is approximately 100 m beneath the surface. The LHC was designed to collide two proton beams at 14 TeV center-of-mass energy.

The LHC started operation in March 2010 at 7 TeV, and ran at 8 TeV during the year 2012. This thesis is based on the 8 TeV period. Although currently shut down, the LHC is expected to operate at a center-of-mass energy up to 13 or 14 TeV when it turns on again in 2015.

The LHC sends two beams of proton bunches in opposite directions, and collides them at four points around the ring. One of these collision points, interaction point five (near Cessy, France), is where the CMS detector is located. There are 3564 bunches per orbit, and the bunches cross at a rate of 40 MHz.

2.1.1 Detector Coordinate System

It is useful to define a coordinate system to describe the position of particles in the CMS detector. The origin of the coordinate system is taken to be the collision point. The positive x axis points towards the center of the ring, the positive y axis points vertically upwards, and the positive z axis is in the direction of the counterclockwise beam as viewed from above. The azimuthal angle (ϕ) is measured from the x axis in the x-y (transverse) plane, and the polar angle (θ) is measured from the z axis in the r-z plane, where r is the radial coordinate in the transverse plane.

The center-of-mass frame for particles produced in a hard-scattering process in proton-proton collisions is boosted longitudinally (i.e., along the beam direction), because only a single quark or gluon constituent of the proton participates in the collision and not the entire proton. It is therefore convenient to use transverse quantities, such as transverse

momentum (p_{T}) or transverse energy ($E{\cdot}\sin\theta$ or E_T), because such quantities are invariant under longitudinal boosts. Additionally, instead of using the coordinate θ, the rapidity (y) is used, which is Lorentz invariant under a longitudinal boost. y is defined as $\frac{1}{2}\ln(\frac{E+P_L}{E-P_L})$. For particles with a small mass, where $E \sim P$, the rapidity is approximated by the pseudo-rapidity η, which is a function of θ only:

$$y \approx \eta \equiv -\ln\tan\frac{\theta}{2} \tag{2.1}$$

Additionally, a useful variable for determining the distance between objects is the space angle Δ_R, which is defined as $\Delta_R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

2.2 Compact Muon Solenoid Detector

CMS is a general purpose detector whose goals include probing physics at the TeV scale, searching for supersymmetry and extra dimensions, and improving precision on measurements of properties of previously discovered particles. CMS was designed using detection of the standard model Higgs boson [6] as a benchmark, which can be considered the chief design goal. The lower mass limit of the Higgs boson in the early 2000s was set by LEP, and corresponded to a lower bound mass of $114\,\mathrm{GeV}$ [42]. Though hadronic decays are the dominant decay mode of the Higgs boson around the mass interval $114 - 130\,\mathrm{GeV}$ [15], the presence of large QCD backgrounds drown out this channel. Final states involving leptons and photons provide a cleaner signature even though the event rate is less than for

the hadronic decay mode. Therefore, clearly identifying leptons and photons was a key consideration in the design of the CMS detector.

A schematic diagram of the CMS Detector is shown in Figure 2.1.

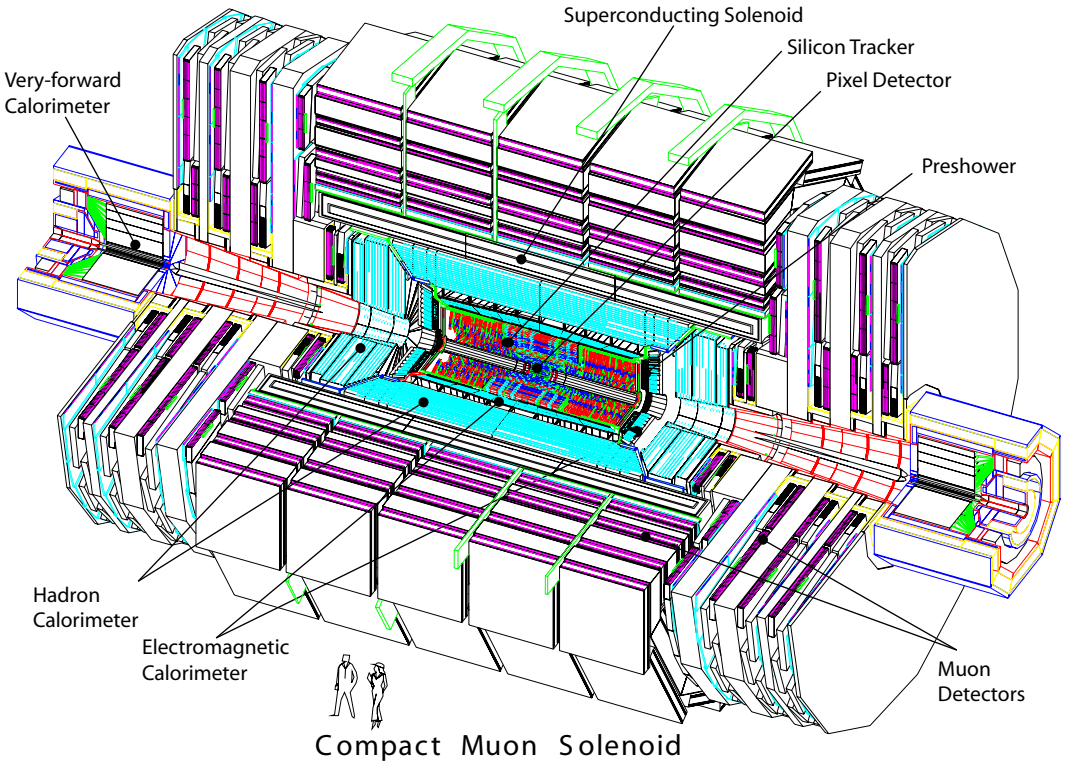


Figure 2.1: A view of the CMS detector including subdetectors labels. People are shown to express the size of the detector. From Ref. [6].

The interaction point is located at the center of the detector, and is where the proton-proton collisions occur. CMS is split into "subdetectors," each playing an important role in the reconstruction of an event. CMS covers almost 4π in solid angle around the interaction point, allowing it to measure the global transerve momentum imbalance. Enough granularity is incorporated to determine up to 20 simultaneous collisions between proton bunches.

A more detailed descrition of CMS and its subdetectors can be found in Ref. [6], but a brief description of the subdetectors follows.

2.2.1 Inner Tracker

The inner tracker includes the silicon tracker and pixel detector. Triggering (Section 2.2.6) and offline identification of τ leptons and b-quark jets (Section 3.4.7) require an efficient and precise tracking system. Identifying b-quark jets relies on determining displaced vertices, so pixel detectors in the tracker are required to be close to the interaction region. Thus, the tracker is the closest subdetector to the particle collision. More specifically, the tracker surrounds the interaction point with length 5.8 m, diameter 2.5 m, and covers η from -2.5 to 2.5.

The tracker is designed to have good charged-particle momentum resolution ($1-5\%$) to better measure the curvature of tracks left by charged particles. To limit multiple Coloumb scattering and bremsstrahlung, minimal material is desired. The material chosen must have high granularity, speed, and a lifetime of about ten years. Furthermore, since the tracker receives the highest volume of particles, it must be radiation resistant. To this end, silicon detector technology was chosen and the tracker is made entirely of silicon. The are 66 million pixel channels and 9.6 million strip channels [12].

2.2.2 Electromagnetic Calorimeter (ECAL)

Surrounding the inner tracker is the electromagnetic calorimeter (ECAL). The driving criterion for the ECAL design is to detect the $Higgs \rightarrow \gamma\gamma$ decay, in addition to providing accurate identification of electrons. The ECAL has wide geometric coverage and is designed to give sufficient diphoton and dielectron mass resolution (about 1% at 100 GeV).

Since π^0 mesons are copiously produced in proton-proton collisions, and since π^0 mesons decay almost exclusively to two photon final states, π^0s form the principal background to prompt photon production. To help distinguish prompt photons from $\pi^0 \rightarrow \gamma\gamma$ decay, presampler detectors constructed from lead-scintillator sandwiches are placed before the ECAL detector in the endcap region [5].

Final state photons and leptons are required to be isolated in order to distinguish them from photons and leptons that form during hadronization. Hadronization is the process of forming hadrons (particles made up of quarks) out of free quarks and gluons after a high energy collision. In order to achieve high density and to improve photon and lepton isolation, lead tungstate crystals are chosen for compactness. There are 61,200 crystals in the barrel region, and 7,324 crystals in the endcap. A 5x5 configuration of crystals are arranged to form a "tower," and these towers cover pseudorapidity from -3.0 to 3.0. The towers are in a 5x5 segmentation to match the HCAL segmentation (see Section 2.2.3). Since each crystal has a thickness of 25 radiation lengths, electrons and photons are very likely to deposit their entire energy in the crystal after entering it.

The ECAL measures energy deposition through charged particles. Energy deposition is converted to light proportional to the particle's energy. The light is kept in the crystal through total internal reflection, and is captured by photomultiplier tubes or silicon photodiodes. The light is then converted to an electrical signal that is amplified and sent to the front-end electronics.

2.2.3 Hadronic Calorimeter (HCAL)

The hadronic calorimeter (HCAL) surrounds the ECAL, and measures the energy of hadrons. Additionally, the HCAL indirectly detects weakly-interacting particles through the $E_{\mathrm{T}}^{\mathrm{miss}}$ quantity 3.4.6. Since $E_{\mathrm{T}}^{\mathrm{miss}}$ is the measure of global transerve momentum imbalance, the HCAL must be hermetic and requires large geometric converage and fine segmentation.

HCAL is separated into four regions: the barrel (HB), outer (H0), endcap (HE), and forward (HF) regions. There are 70,000 plastic scintillating tiles and brass absorbers in 16 layers that cover η from -3.0 to 3.0 in the barrel (HB), outer (HO), and endcap (HE) regions. HB covers η for $0.0 < |\eta| < 1.4$, HO covers η for $0.0 < |\eta| < 1.3$, HE covers η for $1.3 < |\eta| < 3.0$, and HF covers $3.0 < |\eta| < 5.0$ The scintillators are segmented in towers of area $\Delta\eta$ x $\Delta\phi$ = 0.087 x 0.087 for $|\eta| < 1.6$, and 0.17 x 0.17 for $|\eta| > 1.6$. The depth of HCAL exceeds up to nine nuclear interaction lengths, which can contain up to 99% of hadronic cascades.

HO, which covers the region $-1.3 < \eta < 1.3$, lies outside the magnet (see Section 2.2.4) and helps to identify late showers and to separate charged pions and kaons from muons.

Signals in HF are generated by Cherenkov light in quartz fibers and read by photomultipliers. HF also contains luminosity monitors to help determine the integrated luminosity.

Because of the importance of determining missing energy, the layers of HCAL are staggered so that particles that traverse gaps between adjacent cells will not be undetected.

The layers alternate between brass absorbers and plastic scintillators. When the particles hit the absorbers, interactions can occur and secondary particles are produced. Light is summed over many layers of tiles, creating a tower. When particles pass through the plastic, a light pulse is created and passed to the readout system described below.

HCAL Readout and Data Monitoring

Readout of data from HCAL is a pipelined system. The data processing is connected in series and the output of one portion is the input of the next one. The analog signals from the calorimeter cells are converted to digital format synchronously with the 40 MHz collisions, and are received from the CMS HCAL Front-end Electronics (FEE), which is located on the HCAL detector. The entire stream of digitized 25 ns time samples are delivered through optical fibers at a rate of 1.6 Gb/s. The 25 ns corresponds to the bunch spacing, and there are 3564 bunches per orbit. The data is first received by the Versa Module Europa (VME) cards, where the HCAL Trigger Readout (HTR) modules are located. The fibers transfer

data at a rate of 1.6 GB/s. These modules form Trigger Primitives (TP) for each calorimeter tower on each bunch crossing. The TP are all transmitted synchronously to the Regional Calorimeter Trigger (RCT). If the TP passes for all subdetectors, a Level 1 Accept (L1A) trigger is generated for that event and is distributed to the data acquisition (DAQ) system.

One challenge of the readout system is to ensure signals throughout the detector are sampled synchronously, and that the digital data are transferred with constant latency to the trigger and DAQ systems. The HCAL readout synchronization is achieved through a series of modules. The Readout Modules (RM) on Readout Boxes (RBX) receive light from HCAL scintillator tile fibers. This light is converted to electrical signals with the Hybrid Photo Diodes (HPD) and is digitized through Charge Integrating Electronic cards (QIE). The QIEs are made up of 4 capacitors called CapIds. Each capacitor integrates over each beam crossing (BX), or time slice. After a time slice is completed, the QIE switches to a subsequent CapId, while the charge on the previous one is collected and digitized. The cells along a tower are added together and sent to the HTR. The HTR contains a pipeline of values collected one-by-one from the CapIds, arranged in a circular configuration, that are read continuously. The RBX contains a Clock Control and Monitoring Module (CCM) to receive clock commands, and a Calibration Module (CM) to transfer laser and LED signals to each HPD pixel. On the HCAL front end, there are 132 RBXs, each with 4 RMs. The CCM distributes clock and control commands to the TTC Receiver (TTCrx) chip. This chip decodes the clock commands and shifts the clock phase in 240 steps of 104 ps to cover the entire 25 ns cycle. Periodic phase scans, about once a month, on the RBX can provide a

way to ensure that the hardware is working properly. This is done by measuring the phase information on each RBX with an oscilloscope. The phase scan is compared to previous phase scans by looking at difference histograms. It is then possible to determine which CCMs are behaving poorly, and to find all defective RBXs.

Corrupt data can occur, and a group called the "HCal Prompt Task Force" [20], of which I was a member of, was formed to address such issues in a timely manner. Often times, pinpointing the solution involves performing raw-data-HEX dumps of an event. For example, during the beginning of September 2009, invalid data (about 1 in 100,000 events) was observed to affect the entire HCAL detector. From a raw HEX dump of an event with invalid data, it was seen that this data occurred at bunch crossing 0, 1, or 2 and corresponded to the QIE reset. The QIE resets are issued every 103 orbits, and at the end of the orbit. When a QIE is reset, non-standard (invalid) data are sent out by the FEE. Furthermore, when an L1A trigger is issued at the very first BXs in an orbit, the first few time samples contain invalid data. However, this is expected and accounted for in the firmware. The raw HEX dump confirmed that the processing time in the firmware was too slow, and the last invalid data due to the QIE reset was actually being picked up in bunch crossing 3562, which is three bunch crossings below 0. The action taken to resolve the problem was to move the QIE reset back by at least 3 bunch crossings to avoid the corrupted data.

2.2.4 Superconducting Solenoid Magnet

Just outside the HCAL barrel lies the superconducting magnet, to which the "S" in "CMS" refers. In a magnetic field, charged particles move in a curved path. Since the particles resulting from a high-energy collision can have large momenta, a strong magnetic field is required to bend the path of fast-moving particles. The magnetic field of CMS is about 3.8 Tesla, and is produced by an 18 kA current flowing through a solenoid of niobium-titanium coils. The solenoid has a radius of 3 m and length 12.5 m. The solenoid magnet coils are held at a temperature of 4.5 K so that they are superconducting, allowing current to flow without resistance.

A flux-return yoke guides the magnetic field and is interleaved within the muon detectors described in the next section. The return yoke is made of steel. Only muons and weakly-interacting particles are expected to pass through this part of the detector due to the absorption of other particles in the inner detectors.

2.2.5 Muon System

The muon detection system is the outermost subdetector of CMS. Muon detection is a powerful tool for separating interesting signal events from high-cross-section background events. One of the cleanest signatures is the $Higgs \rightarrow \mathrm{ZZ} \rightarrow \mu\mu\mu\mu$ decay with four muons in the final state. Muons penetrate all other subdetectors and are finally detected by the muon "chambers" at the edge of CMS, and besides weakly-interacting particles, are likely to be the only particles to make it that far.

The position of a muon is measured by combining tracker information and position information from the muon stations, and the momentum is found by measuring the curvature of the path.

There are 1400 muon chambers. The 250 drift tubes (DTs) in the endcap and 540 cathode strip chambers (CSCs) in the barrel measure the position of muons as they traverse the detector. The 610 resistive plate chambers (RPCs) offer trigger redundancy, and help to reject background noise. RPCs are installed in both the barrel and endcap regions, and consist of two parallel plates, which are an anode and a cathode.

The DTs are 4-cm-wide tubes and contain a stretched wire in a volume of gas. As muons pass through, electrons are knocked off gas atoms and travel towards the positively charged wire through an electric field. The muon position can be reconstructed by determining the path of the electron. The DT chambers each consists of 12 aluminum layers, and there are up to 60 tubes.

In the endcap region, the magnetic field is uneven and the volume of particles is high. CSCs are used here, and are arrays of positively charged anode wires crossed with negatively charged copper strips in a gas volume. Passing muons knock electrons off gas atoms and cause an "avalanche" of electrons. Positive ions travel towards the copper cathode and induce a charge pulse in the strips. Since the anode wires and cathode strips are perpendicular, two coordinates for position are obtained.

2.2.6 Trigger and Data Acquisition

The experiment produces forty million collisions per second. This collision rate would require over 40 gigabytes of data storage per second. If all data were stored, this would amount to about 1.2 exabytes in one year. CMS is estimated to store five petabytes of data annually, so data must be selectively collected. This is achieved through triggers that are fired for selected events.

Proton-proton bunches collide every 50 ns and the detector system must be able to distinguish particles from different events. This is achieved through synchronization of all the electronic readout channels throughout the detector, which allows accurate trigger measurements.

There are two main layers of triggering. The "L1" level is the first level, and is performed online - before the data are stored permanently. Buffers temporarily store the data and a small amount of information is used to decide whether an event should be kept. This information includes the energy of jets, muons, and the magnitude of $E_{\mathrm{T}}^{\mathrm{miss}}$. The online event reconstruction, completed around 1 μs, is performed on fast reprogrammable field-programmable gate arrays (FPGA). About 100,000 in 1 billion events are saved after the L1 trigger.

If events pass through the L1 trigger, data are sent to computers to be analyzed. The second trigger level is the High Level Triggering (HLT) layer. At this point, the event rate is lower and more time is allowed to review the entire event reconstruction and make a triggering decision. The HLT layer reduces the data to allow it to finally be stored on tape,

and the reduction is roughly an order of magnitude in reduction of events. Thus, about 10,000 in 100,000 events are saved after the HLT layer.

2.2.7 Computing and Software

Even after triggering, there are vast amounts of data that need to be accessed by many hundreds of researchers. The data are distributed worldwide to be stored and accessed for data analysis. The system is called the Worldwide LHC Computing Grid (WLCG), and is comprised of four "Tiers." Tier-0 is at CERN, and is where the data is first held in its raw format. Tier 0 also holds the data where the events are reconstructed (RECO format). After a backup is made, the data travels to Tier-1 centers located in seven locations around the world, and the event reconstruction information is used to analyze the event. Tier-1 centers are mainly responsible for creating the re-processing of the raw data into RECO format, and for distributing these data to other Tier-1 centers. Tier-2 and Tier-3 centers hold compressed versions of the data at the processed level (AODSIM), and contain even more centers to localize the data for analysts. While Tier-2 centers help generate Monte Carlo samples and hold calibration files, Tier-3 centers typically contain only analysis files and are used by analysts at the local institution where the Tier-3 center is located. There are about seven Tier-2 centers in the United States, and around 30 Tier-2 centers worldwide. Tier-3 centers are expected to help parallelize computing for analysts from different institutions and are supported by the user community. The computing cluster at UCR is a Tier-3 center.

The CMS collaboration chose the C++ programming language to create a common framework for analysts to use, called the CMS software (CMSSW). All users who analyze data on CMS use CMSSW, with an open-sourced data analysis package called ROOT [60].

Chapter 3

Event Selection

3.1 Overview

Hadronic events are those that contain no leptons in the event. A search for SUSY in hadronic events with b quarks requires the selection of events with multiple energetic jets, large missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, no leptons, and at least one b-quark jet. The signature of interest has large total visible transverse energy, H_T, which is defined as the scalar sum of the transverse momentum of all jets in an event. The large values of H_T are because of the high level of jet activity expected in all-hadronic SUSY events. The following sections describe the data and samples, and physics objects used in the analysis.

3.2 Samples

3.2.1 Data

As Fig. 3.1 shows, the total integrated luminosity in 2012 delivered by the LHC is 23.30 fb^{-1}, and recorded by the CMS detector is 21.79 fb^{-1}. The difference in integrated luminosities between recorded and delivered exists because the trigger is not 100% efficient and because of detector downtime to address malfunctions.

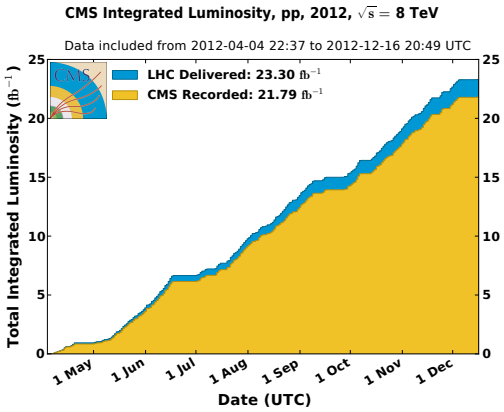


Figure 3.1: Cumulative integrated luminosity in 2012 delivered by the LHC (blue) and recorded by CMS (orange) as a function of day. From [24].

A running cycle, or run, occurs when there are so-called "stable beams" at the LHC. A typical run lasts on average 10 hours (with the longest run being recorded at roughly 23 hours). These are followed by an approximately two hour filling time. A collection of runs is known as a "run era." Different run eras are separated by a gap of a week or more when the LHC is not operated for physics purposes but instead undergoes maintenance or machine development. Different trigger menus exist for different run eras (see Section 3.7.1).

Data are examined, run-by-run, by a data certification team consisting of participants from subdetector and physics object groups. Among other quality checks, such as subdetector performance consistency and stable beams from LHC, histograms of a subset of collected data are reviewed against previous runs. For example, in the "JetMET" group (of which I was a member), whose task is to validate and maintain the performance of jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, histograms of jet and $E_{\mathrm{T}}^{\mathrm{miss}}$ variables are reviewed for every run and compared against past validated runs. For jet objects in particular, variables include the leading and subleading jet p_{T}, η, and ϕ. Certified run ranges are given as files to analysts in JavaScript Object Notation (JSON) format. JSON is a text file format that holds data in a standardized, structured way.

Though the recorded total integrated luminosity for CMS is $21.79\,\mathrm{fb}^{-1}$, only certified runs are used. Several run eras exist within 2012, and are shown with their corresponding certified integrated luminosities in Table 3.1.

Table 3.1: Run eras and their corresponding certified integrated luminosities during the 2012 period.

Run era	$\int \mathcal{L}\, dt$ (pb^{-1})
Run2012A	807
Run2012B	4421
Run2012C	6897
Run2012D	7274
Total	19399

The data considered in the analysis are collected with the CMS detector in 2012, and correspond to $19.4\,\mathrm{fb}^{-1}$ of certified proton-proton collisions at $\sqrt{s} = 8\,\mathrm{TeV}$.

The datasets used in the analysis are chosen based on the trigger choices, which are discussed in detail in Section 3.7.1, and are listed in Table 3.2. The HT/HTMHT, MET and JetHT datasets are used to select the signal and control samples (discussed in Chapter 4), and the DoubleMu and DoubleElectron datasets are used to evaluate the $Z \rightarrow \nu\bar{\nu}$ background (discussed in Section 4.1).

Table 3.2: Data sets used in the analysis.

Dataset	Run range	$\int \mathcal{L}\, dt$ (pb^{-1})
/HT/Run2012A-13Jul2012-v1/AOD	190456-193621	807
/HTMHT/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/HTMHT/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/HTMHT/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/HTMHT/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/MET/Run2012A-13Jul2012-v1/AOD	190456-193621	807
/MET/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/MET/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/MET/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/MET/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/JetHT/Run2012A-13Jul2012-v1/AOD	190456-193621	807
/JetHT/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/JetHT/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/JetHT/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/JetHT/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/DoubleMu/Run2012A-13Jul2012-v1/AOD	190456-193621	800
/DoubleMu/Run2012B-13Jul2012-v4/AOD	193833-196531	4421
/DoubleMu/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/DoubleMu/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/DoubleMu/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/DoubleElectron/Run2012A-13Jul2012-v1/AOD	190456-193621	800
/DoubleElectron/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/DoubleElectron/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/DoubleElectron/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/DoubleElectron/Run2012D-PromptReco-v1/AOD	203768-207469	7274
	Total	19399

3.2.2 Monte Carlo Simulation

Simulated data are essential to this analysis for several reasons. Simulated data allow SM physics processes to be studied individually. If these processes are backgrounds to the experimental signature, they can be investigated independently or collectively to optimize signal selection criteria. Furthermore, modeling physics beyond the standard model is useful for "new physics" searches, and provides a framework for interpreting results.

Monte Carlo methods are applied for situations where the probabilities of a process are known, but determining the results are difficult. These techniques use random numbers to output values that form a probability distribution. The probability distribution is determined from underlying theories. Due to the probabilistic nature of the production and decay of particles in high energy collisions, Monte Carlo techniques are used to generate simulated physics processes. Additionally, the passage of particles through a detector can involve probabilistic processes, such as ionization and Coloumb scattering [62]. Thus, Monte Carlo methods are also used to model the detector response.

The main standard model backgrounds to this analysis are events with a top-quark pair ($t\bar{t}$), events with a single-top quark, and events with either a W or Z boson produced in association with jets (W+jets and Z+jets events). Smaller background event classes contain two vector bosons (diboson events, i.e., events with a WW, WZ, or ZZ pair) or a Drell-Yan process in association with jets. The standard model backgrounds are discussed in Chapter 4. The hard-scattering at the parton level is modeled by PYTHIA [63], MADGRAPH, or POWHEG and uses CTEQ6 [56] for the parton distribution functions. PYTHIA

is used for the subsequent parton showering, hadronization, and decay. The interactions and energy deposits in the detector are simulated with GEANT4 [61]. $t\bar{t}$, W+jets, and Z+jets events are generated with MADGRAPH, single-top events are generated with POWHEG, and PYTHIA is used for QCD and diboson events.

The Monte Carlo (MC) samples are shown in Table 3.3 along with their cross sections and generated integrated luminosity.

Table 3.3: Standard model MC samples used in the analysis, their cross section, and their generated integrated luminosity.

MC SM Process	σ (pb)	$\int \mathcal{L}\, dt$ (fb^{-1})
QCD Multijet	varies (LO)	
$t\bar{t}$ + jets	234 (NNLO)	29.6
$t\bar{t}$ full leptonic decays	26 (NNLO)	462.9
$t\bar{t}$ semi leptonic decays	104 (NNLO)	245.1
$t\bar{t}$ hadronic decays	104 (NNLO)	299.9
Single Top, s channel	3.79 (NLO)	68.6
Single anti Top, s channel	1.76 (NLO)	79.5
Single Top, t channel	56.4 (NLO)	0.42
Single anti Top, t channel	30.7 (NLO)	63.0
Single Top, tW channel	11.1 (NLO)	44.8
Single anti Top, tW channel	11.1 (NLO)	44.5
W+jets $250\,\text{GeV} < H_T < 300\,\text{GeV}$	57.3 (NNLO)	86.2
W+jets $300\,\text{GeV} < H_T < 400\,\text{GeV}$	45.7 (NNLO)	112.5
W+jets $H_T > 400\,\text{GeV}$	30.1 (NNLO)	165.2
Z $\to \nu\bar{\nu}$ $100\,\text{GeV} < H_T < 200\,\text{GeV}$	205.2 (NNLO)	52.4
Z $\to \nu\bar{\nu}$ $200\,\text{GeV} < H_T < 400\,\text{GeV}$	53.1 (NNLO)	197.4
Z $\to \nu\bar{\nu}$ $H_T > 400\,\text{GeV}$	6.75 (NNLO)	812
Drell-Yan $200\,\text{GeV} < H_T < 400\,\text{GeV}$	23.43 (NLO)	295
Drell-Yan $H_T > 400\,\text{GeV}$	3.36 (NLO)	812
WW	55 (NLO)	182
WZ	32.3 (NLO)	310
ZZ	17.654 (NLO)	555
γ+jets $H_T > 400\,\text{GeV}$	107.5 (NNLO)	15.0

Several H_T-binned samples create the QCD MC dataset, and this is chosen to increase the size of the event sample in the higher H_T regions. Otherwise, the events will be dominated by low H_T QCD events due to the higher cross sections there. The QCD distributions throughout this thesis are normalized to the LO cross sections, and the single-top, diboson, and Drell-Yan distributions are normalized to the NLO cross sections. The $t\bar{t}$, W+jets, Z+jets, Z $\rightarrow$ $\nu\bar{\nu}$, and γ+jets distributions are normalized to the NNLO cross-sections. The γ+jets events are used to explore a method to evaluate the Z+jets background, as described in Section 4.1.1.

The standard model cross section factors are obtained from Ref. [32], and are dependent upon the physics process. When the Monte Carlo sample for a particular process is created, the cross section is fixed and the number of events requested is generated. With higher number of events comes better statistical precision, but storage space and processing time are limited resources that must be considered. To match a given integrated luminosity in data, the Monte Carlo sample must be reweighted. This weight is given by the following formula:

$$weight = \frac{\mathcal{L}_d \cdot \sigma(pb)}{N_{events}} \tag{3.1}$$

where $\mathcal{L}_d$ is the integrated luminosity of data being compared, σ (pb) is the cross section of the Monte Carlo sample, and N_{events} is the number of generated Monte Carlo events.

The MC signal samples T1bbbb and T1tttt (see Section 3.3) are generated with MADGRAPH over a range of m_{gluino} and LSP m_{LSP} mass values. These samples are modeled

with the CMS FASTSIM program [22], which is chosen to reduce computational resources. Modest differences, such as for the b-jet tagging efficiency, are found between the GEANT4 and FASTSIM simulation, and a correction is applied to account for this difference.

3.3 Simplied Models with b quarks

In order to model a SUSY signal, simplified models (SMS) are considered. An SMS is not a full-fledged SUSY model, but is a more simplified version. It is assumed that a certain non-standard model process, represented by an SMS, is the only one observed in data. In this thesis, the signal process is represented by either the so-called "T1bbbb" or "T1tttt" model. A diagram of these models is shown in Fig. 3.2. T1bbbb (T1tttt) is described by gluino pair production followed by the decay of each gluino $\tilde{g}$ to a b quark and an off-shell b squark. The off-shell b squark decays into a b (t) quark and the LSP, which is assumed to be the lightest neutralino $\tilde{\chi}_1^0$. The LSP escapes detection and leaves a trail of significant $E_{\mathrm{T}}^{\mathrm{miss}}$ (see Section 3.4.6).

The LSP is assumed to be stable in the SMS. For T1bbbb and T1tttt, the gluino mass (m_{gluino}) and LSP mass (m_{LSP}) are free parameters, and are varied in a scan, with $m_{\mathrm{LSP}} < m_{\mathrm{gluino}}$. Using the MADGRAPH MC generator, around 130,000 events are simulated for each mass point. The analysis can then exclude points at the 95% confidence level if no discovery is found, using a provided reference cross section for gluino pair production.

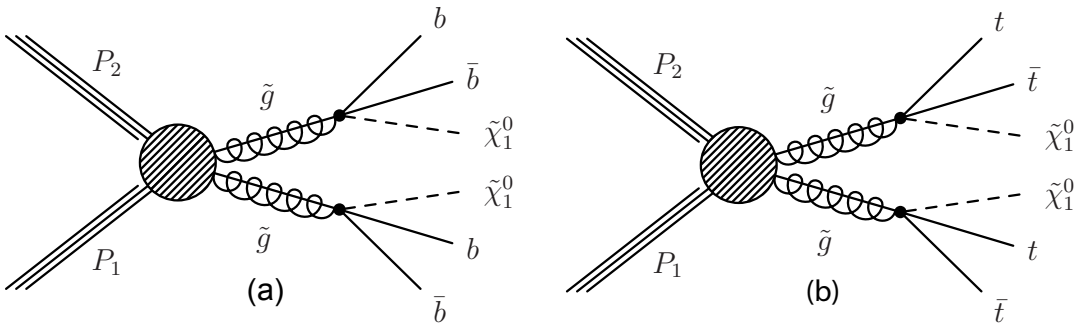


Figure 3.2: Event diagram for the (a) T1bbbb and (b) T1tttt simplified model new-physics scenarios.

The SMS simplifies analyses because we only have to look for events with multiple b-quark jets (see Section 3.4.7) and significant momentum imbalance. The results of the analysis can then be applied to many different SUSY models that have these particular signatures.

Throughout this thesis, the T1bbbb or T1tttt model may also be referred to as "SUSY" or "signal."

3.4 Physics Objects

Chapter 1 discussed the standard model and briefly what quarks and leptons are, but these particles leave certain signatures at the detector level. A discussion of all the physics objects used in this thesis follows.

3.4.1 Primary Vertex

A primary vertex (PV) represents the reconstructed position of the proton-proton collision of interest in an event. The PV is defined as the reconstructed vertex with the highest value of summed ${p_{\mathrm{T}}}^2$ in the event, where the p_{T} is the transverse momentum of charged tracks associated with the vertex. Tracks are grouped into clusters based on their proximity to the beam line. A fit is performed on each cluster of tracks to determine the reconstructed vertex position. To ensure the proximity of the PV to the interaction point, a cut of $|z| < 24$ cm and $\rho < 2$ cm with respect to the origin is required.

3.4.2 Particle Flow algorithm

The basic physics objects, specifically charged and neutral hadrons, electrons, and photons are reconstructed using the so-called Particle Flow (PF) algorithm [21]. The PF algorithm combines information from the all the subdetectors. This information consist of all the stable, or final state, particles produced from the proton-proton collision. In particular, the PF algorithm collects reconstructed hits in all the subdetectors and creates "blocks" that consist of tracks or energy deposits. The blocks are then linked, creating particle flow candidates (PF candidates).

These PF candidates are cross-cleaned according to the SUSYPAT [64] recipe as implemented in Ref. [16]. So-called "cross cleaning" is required to eliminate ambiguities in the event reconstruction. For instance, an energy deposit in the calorimeter might lead to reconstruction of both a photon and a jet. For the purpose of modeling $Z+\mathrm{jets}$ events using

the γ+jets events (see Section 4.1.1), for example, it is important that an isolated photon not be treated as a jet. Thus the jet associated with the photon needs to be removed from the jet collection. The cross-cleaning procedure also eliminates object duplicates from the event.

The types of PF candidates are now discussed.

3.4.3 Jets

Jets represent the final states of quarks and gluons in a decay process and are reconstructed from energy depositions, or towers, in the hadron and electromagnetic calorimeters. These energy deposit towers are clustered by the anti-k_{t} algorithm [14] with a size parameter of 0.5. In the clustering algorithm, each tower is assigned a momentum. The absolute value of the momentum, or raw jet energy, is given by the sum of the energies in the clustered towers, and the direction can be calculated based on the coordinates of the tower. The raw jet energies are then corrected to establish a relative uniform response of the calorimeter in η, and a calibrated absolute response in transverse momentum p_{T}.

The specific clustering algorithm used to create jets is the anti-k_{t} sequential clustering algorithm [14]. This algorithm, which is infrared and collinear safe, uses two distance measures:

- the distance d_{ij} between two particles i and j: $d_{ij} = \min(k_{T_i}^{2p}, k_{T_j}^{2p})\dfrac{\Delta_{ij}}{D}$, where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$

- the distance d_{iB} between any particle i and the beam B: $d_{iB} = {k_{T_i}}^{2p}$

The smallest result is found from among all values of d_{ij} and d_{iB}. If the smallest result comes from d_{ij}, the algorithm combines the two particles *i* and *j* by summing their four-momenta. A new smallest measurement is computed and the process is repeated. If the smallest result comes from d_{iB}, a jet is reconstructed in the direction of particle *i*, and particle *i* is removed from computation. The process is repeated until no particles are left.

The two parameters in the algorithm are D and p:

- D: a scaling factor between d_{ij} and d_{iB} to ensure that any pairs of jets *a* and *b* are separated by at least $\Delta_{ab}^2 = D^2$. The value used in this thesis is D = 0.5.

- p: used to adjust the relative strength of the energy versus geometry scales, and is assigned a value of -2 in the anti-k_{t} algorithm

The anti-k_{t} algorithm produces circular cone-shaped jets relatively insensitive to the so-called "underlying event." The underlying event includes interactions that occur outside of the primary hard scattering interaction, such as initial and final state radiations, beam remnants, or so-called "pile-up." Multiple interactions can occur in the same bunch crossing, and pile-up refers to the uninteresting interactions that occur in the same bunch crossing as an interesting interaction. Energy corrections, which rely on MC truth (generator level) information and data-driven methods, are applied to the cross-cleaned jets by applying a scaling factor to the jet four-momentum for each of the following categories in sequential order:

- offset: subtract unwanted excess energy from noise and pile-up.

- relative: ensure uniformity in jet response versus pseudorapidity by choosing an eta of an arbitrary jet and correcting relative to the central region, where jet response is defined as $< \frac{\text{CaloJet} \quad p_{\text{T}}}{\text{GenJet} \quad p_{\text{T}}} >$.

- absolute: ensure uniformity in jet response versus p_T by correcting back to the particle level.

CaloJets are jets that are reconstructed using calorimeter information only, while GenJets are jets reconstructed in simulated events assuming perfect detector resolution and acceptance.

In the analysis described in this thesis, a jet is required to have:

- $p_{\text{T}} > 50\,\text{GeV}$

- $|\eta| < 2.4$

The last cut requires the jet to be contained within the hadronic calorimeter. Additionally, jets are required to pass the "loose" jet ID criteria [46]:

- neutral EM fraction < 0.99,

- number of constituents > 1,

- charged hadron fraction > 0,

- charged multiplicity > 0,

- charged EM fraction < 0.99

These "loose" jet ID criteria are used to reject fake jets due arising from calorimeter noise. The definition of each jet ID criterion is described in detail in Ref. [1].

Jet validation

Jet algorithms are ever evolving, for example, due to changes in the jet energy correction implementation, so a necessary task involves maintaining the jet reconstruction code. The main duties involve checking the software to ensure that the contributed jet reconstruction code compiles without errors, and that the dependent packages are included in each software release cycle. Additionally, tests are made on locally prepared samples, and the changes are validated by verifying that they match expectations. After new code passes the validation tests performed by jet reconstruction validation experts, of which I served as one, an announcement is posted on the so-called "hypernews" forum containing a short description.

3.4.4 Leptons

For this analysis, electrons and muons are explicitly reconstructed, while tau leptons are not. The selection criteria for electrons and muons are now discussed.

Electrons

A PF candidate is an electron if the PF algorithm creates a link between a charged track from the tracker and ECAL clusters, in addition to following a set of identification criteria. These criteria are influenced by the Physics Object Group (POG) recommendation for what defines a PF electron object. The selection in this thesis follows the "Cut Based VETO" selection (see Ref. [33]), which requires the electron to have $p_{\mathrm{T}} > 10\,\mathrm{GeV}$, $|\eta| < 2.5$, and be isolated. In order for an electron to be isolated, it must have $\sigma_{i\eta i\eta} < 0.01(0.03)$ in the barrel (endcap) region, which is related to the η width of the cluster of hits. The electron candidate transverse impact parameter with respect to the beam spot must not be greater than 0.04 cm, and the longitudinal displacement from the primary vertex must be less than 0.2 cm. Additionally, the hadronic-over-electromagnetic energy fraction must be less than 0.15. The relative isolation, defined as $(IsoTRK + IsoECAL + IsoHCAL)/e_{p_{\mathrm{T}}}$, where $e_{p_{\mathrm{T}}}$ is the electron candidate p_{T}, must be less than 0.15. IsoTRK is the scalar sum of the transverse momenta of reconstructed tracks within a cone of radius $\Delta_R = 0.3$ around the reconstructed muon candidate. IsoECAL and IsoHCAL are similar to IsoTRK, except instead of the transverse momenta, the transverse energy measured in the electromagnetic and hadron calorimeters, respectively, are used.

Muons

Muons are determined from PF candidates by creating links between isolated tracks from the tracker, ECAL clusters, and tracks in the muon subdetector. Muons are required to have

$p_{\mathrm{T}} > 10\,\mathrm{GeV}$, $|\eta| < 2.4$, and be isolated. In order for a muon to be isolated, it must be identified as a global and tracker muon, where at least one muon chamber hit is used in the global fit that matches muon segments in at least two muon stations. Furthermore, the number of valid tracker hits must be at least six, where one of the hits is a pixel hit. The transverse impact parameter with respect to the beam spot must not be greater than 0.2 cm, and the longitudinal displacement from the primary vertex must be less than 0.5 cm. The relative isolation, defined as $(IsoTRK + IsoECAL + IsoHCAL)/\mu_{p_{\mathrm{T}}}$, where $\mu_{p_{\mathrm{T}}}$ is the muon candidate p_{T}, must be less than 0.2. The isolation variables must have reconstructed tracks within a cone of radius $\Delta_R = 0.4$ around the reconstructed muon candidate. The selection follows closely the POG-recommended "Tight Muon" selection (see Ref. [34]).

3.4.5 Total Visible Transverse Energy

The total visible transverse energy, H_T, is the scalar sum of the transverse momentum of all jets in an event passing the jet selection described in Section 3.4.3. H_T is useful in characterizing all-hadronic SUSY signal events such as are the topic of this thesis, because such events lead to long decay chains with many jets and large values of H_T.

3.4.6 Missing Transverse Energy

Were all the objects created in a collision accounted for with perfect resolution and acceptance, the sum of the vectorial transverse momentum will be zero due to the conservation of total momentum. However, some particles escape the detector without interaction. Such

particles include neutrinos, the lightest neutralino found in many SUSY models, or other weakly interacting particles not yet discovered. In addition, imperfect reconstruction of jets and other physics objects can lead to non-zero summed transverse momentum.

Missing transverse energy, or $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as the negative vectorial transverse energy sum of all PF objects in the event. $E_{\mathrm{T}}^{\mathrm{miss}}$ represents the undetected final state particles in an event, and is important in searches for new weakly interacting particles. Due to the necessity of calculating the sum of transverse momenta of all objects in the event, $E_{\mathrm{T}}^{\mathrm{miss}}$ is computed only after all other objects have been reconstructed.

$E_{\mathrm{T}}^{\mathrm{miss}}$ can be influenced by pile-up collisions. The measurement can further be affected by the different responses between photons and pions in the ECAL and HCAL detector. Therefore, corrections are applied to the $E_{\mathrm{T}}^{\mathrm{miss}}$ quantity. Type-I corrections are applied, according to Ref. [17], which are propagations of the jet energy corrections (see Section 3.4.3) to $E_{\mathrm{T}}^{\mathrm{miss}}$.

3.4.7 B-tagging

To identify b jets, tagging algorithms have been developed by the CMS b-tagging physics object group (bPOG). Many algorithms exist for identifying b jets. They rely on the displaced vertex of a b jet compared to the primary vertex, large impact parameter of tracks from a b jet, or kinematical properties, such as jet mass, that distinguish b jets from other types of jets. The one chosen for this search is recommended by the b POG. In addition

to the selection criteria defined by the b POG, the b jets must also satisfy the jet criteria outlined in Section 3.4.3.

The combined secondary vertex (CSV) tagger is used, and the "medium" working point is chosen, which is a value of 0.679. This tagger is described in more detail in Ref. [19]. This working point gives a mistag rate of 1% for light-quark jets, and has an efficiency of 75% for b jets with $p_{\mathrm{T}} = 80$ GeV, as determined from a sample of simulated top-antitop quark pair events.

3.5 Event Selection

The following subsections discuss the $\Delta\hat{\phi}_{\mathrm{min}}$ variable and further event cleaning, and the final subsection summarizes all the baseline selection cuts required for the analysis in this thesis.

3.5.1 $\Delta\phi_N$ Variable

In a QCD event, a single badly mismeasured jet can give rise to large $E_{\mathrm{T}}^{\mathrm{miss}}$. Traditionally, a variable $\Delta\phi_{min}$ is used to reject QCD events. $\Delta\phi_{min}$ is determined by computing the difference in the transverse angle, $\Delta\phi$, between each jet and $E_{\mathrm{T}}^{\mathrm{miss}}$, and then finding the minimum of these values. $\Delta\phi_{min}$ provides a discriminate between QCD and non-QCD events, since QCD events typically have a small value of $\Delta\phi_{min}$ compared to non-QCD events. However, this variable has a strong correlation with $E_{\mathrm{T}}^{\mathrm{miss}}$.

A novel variable, $\Delta\hat{\phi}_{\rm min}$, was developed for this analysis, primarily by Professor Owen Long of UC Riverside, to provide rejection of QCD events while having a weak correlation with $E_{\rm T}^{\rm miss}$. $\Delta\hat{\phi}_{\rm min}$ is described in more detail in Refs. [65, 66], but a brief description is given here.

Assuming there is one badly mismeasured jet and all other jets in the event are perfectly measured, $E_{\rm T}^{\rm miss}$ will point exactly in the direction of the mismeasured jet. Realistically, the other jets are typically also mismeasured by an amount approximated by their energy resolutions. We assume that only one jet is badly mismeasured, and that jet mismeasurement is dominated by jet $p_{\rm T}$ mismeasurement, which we assign to be a 10% [37] resolution on the jet energy. This implies that only additional jets, and not jet angular mismeasurement, will affect $\Delta\phi$ between any jet and $E_{\rm T}^{\rm miss}$.

We define a variable, Δ_{Tj}, which is the perpendicular $E_{\rm T}^{\rm miss}$ component to jet j. This variable is computed through the equation

$$\Delta_{Tj} = \sqrt{\sum_{i=1}^{i<n_{\rm jets}} \left[\sigma(p_T^i)\,\sin(\Delta\phi_{ij})\right]^2} \approx 0.1\sqrt{\sum_{i=1}^{i<n_{\rm jets}} \left[p_T^i\,\sin(\Delta\phi_{ij})\right]^2} \qquad (3.2)$$

where $\Delta\phi_{ij}$ is the $\Delta\phi$ value between jet j and jet i. σ is the 10% resolution on the jet energy as stated above, and p_T^i is the $p_{\rm T}$ of jet i. The summation is over all other jets in the event that are not jet j, since when $i = j$, $\sin(\Delta\phi_{ij}) = 0$.

If jet j is the badly mismeasured jet and all other jets are mismeasured only by their jet energy resolutions, then $\sin^{-1}(\Delta_{Tj}/E_{\rm T}^{\rm miss})$ will be approximately $\Delta\phi(jet_j, E_{\rm T}^{\rm miss})$. How-

ever, if jet j is not a badly mismeasured jet but there exists another jet in the event that is badly mismeasured, then $\sin^{-1}(\Delta_{Tj}/E_{\mathrm{T}}^{\mathrm{miss}})$ will underestimate $\Delta\phi(jet_j, E_{\mathrm{T}}^{\mathrm{miss}})$, since the value of Δ_{Tj} will be underestimated due to incorrectly assigning the mismeasurement of the badly mismeasured jet equal to the jet energy resolution.

We define the normalized $\Delta\phi$ between jet j and $E_{\mathrm{T}}^{\mathrm{miss}}$ as

$$\Delta\phi_{Nj} = \Delta\phi(jet_j, E_{\mathrm{T}}^{\mathrm{miss}})/\sin^{-1}(\Delta_{Tj}/E_{\mathrm{T}}^{\mathrm{miss}}), \tag{3.3}$$

This value is computed for the three highest-p_{T} jets, and we call the minimum of these values $\Delta\hat{\phi}_{\mathrm{min}}$.

In a QCD event containing one badly mismeasured jet, $\Delta\hat{\phi}_{\mathrm{min}}$ will be close to unity, since the smallest value of $\Delta\phi_N$ will be close to unity when it is computed for the badly mismeasured jet. All other jets in this event will have a $\Delta\phi_N$ greater than unity, if they are not badly mismeasured.

For events containing a neutrino that produce genuine $E_{\mathrm{T}}^{\mathrm{miss}}$, such as $t\bar{t}$, $W+\mathrm{jets}$, and single-top events, $\Delta\hat{\phi}_{\mathrm{min}}$ will be greater than unity since the neutrino will not tend to align with any jet in the event.

Because $\Delta\phi(jet_j, E_{\mathrm{T}}^{\mathrm{miss}})$ and $\sin^{-1}(\Delta_{Tj}/E_{\mathrm{T}}^{\mathrm{miss}})$ both scale with $E_{\mathrm{T}}^{\mathrm{miss}}$, the ratio of these values gives a more independent correlation with $E_{\mathrm{T}}^{\mathrm{miss}}$. The distributions for $\Delta\phi_{min}$ and $\Delta\hat{\phi}_{\mathrm{min}}$ over several $E_{\mathrm{T}}^{\mathrm{miss}}$ bins in a QCD MC sample are shown in Fig. 3.3. It is seen that while $\Delta\phi_{min}$ has a strong correlation with $E_{\mathrm{T}}^{\mathrm{miss}}$, $\Delta\hat{\phi}_{\mathrm{min}}$ has a weak correlation with $E_{\mathrm{T}}^{\mathrm{miss}}$.

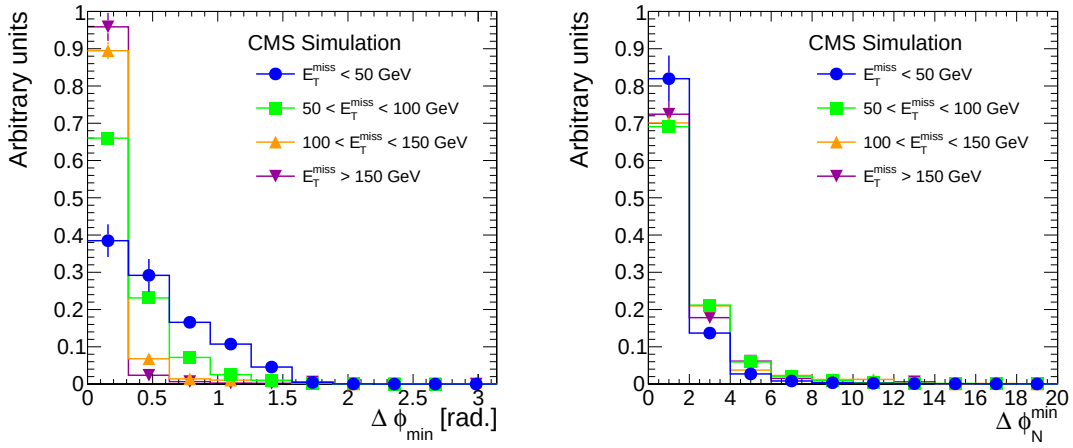


Figure 3.3: Distributions of (left) $\Delta\phi_{min}$ and (right) $\Delta\hat{\phi}_{\mathrm{min}}$ in bins of $E_{\mathrm{T}}^{\mathrm{miss}}$ in MC QCD events with at least 1 b-tagged jet. From [65].

3.5.2 $E_{\mathrm{T}}^{\mathrm{miss}}$ Cleaning

Anomalous events can arise due to event misreconstruction or beam noise. This can cause "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$ in the detector, since they can contribute momentum, separate from the physics process in the interaction, that thus causes mismeasured momentum imbalance. These events can be "cleaned", or rejected, through various filters. Anomalous events, especially in the tail of the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution, can cause significant implications for a signal discovery. It is important to remove events that are traced back to detector sources that mimic $E_{\mathrm{T}}^{\mathrm{miss}}$ [47]. Several filters are applied to reduce or eliminate such anomalous events. These filters are designed to reject events with large fake $E_{\mathrm{T}}^{\mathrm{miss}}$ but to have negligible impact on signal efficiency. The $E_{\mathrm{T}}^{\mathrm{miss}}$ tail cleaning is discussed in more technical detail in Ref. [66].

3.5.3 Isolated Track Veto

$t\bar{t}$ and $W+\text{jets}$ events that pass the baseline selection (see Section 3.5.4) typically have an unidentified lepton. This lepton can be a tau lepton that subsequently decays hadronically. To reduce background events with these hadronic taus, events are rejected that contain at least one isolated track with $p_{\mathrm{T}} > 15\,\text{GeV}$ and $|\eta| < 2.4$, and with relative isolation less than 0.05 within a cone of 0.3. The relative isolation considers tracks that have p_{T} with at least 2 GeV, and are within $|z| < 0.05cm$ from the PV. The isolated track veto also eliminates background events with an isolated electron or muon that are unidentified.

3.5.4 Baseline Selection

A summary of the baseline selection is listed in this section.

- at least three PF jets, where the leading two jets must have $p_{\mathrm{T}} > 70\,\text{GeV}$ due to the HLT_DiCentralPFJet50_PFMET80 trigger efficiency (see Section 3.7.1);
- $H_T > 400\,\text{GeV}$;
- $E_{\mathrm{T}}^{\mathrm{miss}} > 125\,\text{GeV}$;
- electron and muon veto;
- $\Delta\hat{\phi}_{\mathrm{min}} > 4.0$;
- $E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning;
- isolated track veto.

The electron, muon, and isolated track veto reduces background from top and $W+\text{jets}$ events (where "top" refers both to $t\bar{t}$ and single-top quark events). The $\Delta\hat{\phi}_{\text{min}}$ variable helps reduce QCD background events, as discussed in Section 3.5.1.

3.6 Ntuples

In principle, a CMS analysis could be performed directly on the AODSIM (see Section 2.2.7) files. However, these files are often big, which make them time consuming to run on and challenging to store. To alleviate this issue, ntuples, which are ROOT files that contain only a minimal set of variables required for an analysis, are created. For the current analysis, ntuples are created that store basic information for each event such as $E_{\text{T}}^{\text{miss}}$, H_T, the number of jets, the number of b-jets, the number of electrons, etc., where all the objects have been selected according to the criteria outlined above. Analysis cuts are then imposed on ntuples so that physics distributions, inputs to fits, or event counts are extracted.

It is generally redundant for each analyst to create an independent set of ntuples, due to the numerous common variables. Therefore, analysts working in the same group or on the same topic typically create common ntuples. These ntuples are usually larger than would be required by an individual analyst, but the time saved in generating common ntuples is often worth the size costs. Furthermore, since the ntuples are of smaller size, the time spent running on them is also less. Because the common ntuples are produced for general use,

additional time costs include transferring these common ntuples from a central location to a local location for analysis. However, these time costs are relatively infrequent.

In this analysis, the common ntuples were so-called "cfA ntuples", where "cfA" stands for "Configurable Analysis." We used an interface provided by the UC Santa Barbara group to manage the generation of cfA ntuples. The cfA ntuples were then transferred from the computing center at UC Santa Barbara to local computing areas. In this analysis, I generated several cfA ntuples at UC Santa Barbara, and transferred all of the cfA ntuples to the UC Riverside Tier 3 cluster for use by UCR analysts.

3.6.1 Synchronization Studies

As a means to validate analysis code and help identify programming errors, it is useful for independent researchers working on a common ntuple to compare their results for the numbers of events obtained after each cut in a baseline selection. Any discrepancies indicate differences due to such things as different calibrations or corrections (which can arise as a consequence of misunderstandings between analysts) as well as outright errors. Such an exercise, known as synchronization, is extremely valuable and can be surprisingly complex. Synchronization ensures confidence in correctly applying the selection cuts and verifies the consistency of all the samples used. For the analysis in Ref. [65], four groups synchronized on two independently produced ntuples. These groups include UC Riverside, UC Santa Barbara, Cornell U., and U. Colorado at Boulder. A high level of synchronization is achieved, where the results are shown in Ref. [28].

The synchronization effort from UC Riverside was performed entirely by me. After transferring the common cfA ntuples to the UC Riverside Tier 3 cluster, I created so-called "BigNtuples," which stripped the cfA ntuples to only the necessary variables required for my analysis. These BigNtuples are generated infrequently, and often are generated only once after transferring the common cfA ntuples. In the next stage, I created mini-files, which contain the same variables as BigNtuples, but with the baseline selection already applied. These can reduce a file up to ten times, e.g., from 10 MB to 1 MB, and allows for a reduction in processing time (from several hours to several minutes). The synchronization studies are performed on the BigNtuples, since the mini-files already have the baseline selection applied. It took around four to five rounds be highly synchronized between all the groups.

3.7 Trigger

This section describes the trigger efficiency measurements used in this thesis, and the trigger systematic uncertainties. I played a crucial role in selecting the triggers used in this analysis, measuring trigger performance, and evaluating trigger systematic uncertainties.

3.7.1 Trigger Efficiencies

As discussed in Section 2.2.6, limited bandwidth and storage capacity precludes analysts from recording the result of every collision that occurs within the detector. The trigger is

the system used to determine which events are recorded and which are not. Trigger decisions thus represent a balance between reducing the number of events to a manageable level and the desire to keep as many events of potential interest as possible. This balance can be challenging due to similarities between interesting and uninteresting events, and the restriction of designing a set of triggers that will satisfy all the analysis groups including searches for the Higgs boson, searches for physics beyond the SM, and SM precision measurements. The possible triggers are collected in a trigger "menu." Selecting the most appropriate collection of triggers is a crucial requirement for an analysis striving for a discovery based on a low event rate process, or in the case of no discovery, setting strong limits. This section discusses the formulation of the triggers used in the analysis described in Ref. [31].

Triggers are formed by combining a few selection criteria of physics objects, such as the number of jets with a particular p_{T} value, a particular $E_{\mathrm{T}}^{\mathrm{miss}}$ value, a particular H_T value, etc. For instance, a single trigger can combine several criteria, or use only one. When several criteria are required, the trigger condition is satisfied when all the individual criteria are satisfied (logical 'AND'). These triggers are also called "cross triggers."

The CMS trigger contains a level 1 (L1) compoment, which is based on online-level information as discussed in Section 2.2.6. At the online level, physics objects are created with Calorimeter information (CaloJets, CaloMET), as opposed to PF, because the calo information is available with a much faster response time.

There are two types of triggers: prescaled and unprescaled. Prescaled triggers are those where the thresholds on the selection criteria of the trigger are low enough that the event

rate is still high. The prescaled trigger caries a factor, called a prescale, which is defined as the fraction of events recorded passing the prescaled trigger. Therefore, not all events passing the prescaled trigger are actually accepted. Unprescaled triggers, on the other hand, accept all events and have tighter thresholds in their selection criteria. To increase the number of events, which can reduce statistical uncertainty, the analysis described in this thesis prefers unprescaled triggers. Even though all events are kept in unprescaled triggers, enough space is available since tighter cuts correspond to fewer events, and thus less storage requirements.

Several datasets exist for different physics goals. For instance, analyses that use $E_{\mathrm{T}}^{\mathrm{miss}}$ as a selection criteria might consider the MET dataset, which contains logical OR combination of many $E_{\mathrm{T}}^{\mathrm{miss}}$-based triggers. An analysis that looks at multiple jets might consider the JetHT dataset, which contains a logical OR combinations of many cross triggers on the number of jets and H_T. The datasets used for measuring trigger performance in this analysis are shown in Table 3.4. The name gives a clue to what triggers are used, however the exact triggers contained in the dataset are based on the run period, CMSSW software, and can be found in the web-interface tool in Ref. [26]. Each dataset is composed of a logical OR of a set of triggers, but the information of all triggers in every dataset is stored in each dataset. This information includes whether a trigger would have fired or not for each event, which is important for measuring trigger performance.

An initial choice for a trigger is one that is a subset of the analysis baseline selection described in Section 3.5.4 . In a previous analysis [27], the H_T selection was required to be

at least 350 GeV, and $E_{\mathrm{T}}^{\mathrm{miss}}$ was required to be at least 200 GeV. Therefore, an unprescaled trigger was created specifically for this analysis called HLT_PFHT350_PFMET100. We use this trigger for the 2012 data analysis described here. As the name implies, this trigger requires PF $H_T > 350\,\mathrm{GeV}$ and PF $E_{\mathrm{T}}^{\mathrm{miss}} > 100\,\mathrm{GeV}$. Here, the PF H_T in the PF H_T leg is defined as the transverse momentum sum of all PF Jets with $p_{\mathrm{T}} > 40\,\mathrm{GeV}$ and $|\eta| < 3.0$. The PF $E_{\mathrm{T}}^{\mathrm{miss}}$ leg is defined as in Section 3.4.6, and is also a logical OR with $H_{\mathrm{T}}^{\mathrm{miss}}$ of at least 150 GeV. $H_{\mathrm{T}}^{\mathrm{miss}}$ at the trigger level is the negative vector sum of the transverse momentum of Calo jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$. As explained previously, the Calo component of the trigger allows for a faster response.

The reference trigger is a trigger used to measure the efficiency of an analysis trigger with respect to the offline analysis cuts. An analysis trigger is fully efficient if all the events passing the analysis trigger also pass the offline analysis cuts, and inefficient if events do not pass the trigger but pass the offline analysis cuts. The reference trigger is needed to ensure the efficiency is measured consistently, and the only variable being measured is the analysis trigger. The formula to determine trigger efficiency is defined by:

$$analysis\ trigger_{eff} = \frac{analysis\ trigger \times reference\ trigger \times offline\ cuts}{reference\ trigger \times offline\ cuts} \tag{3.4}$$

The reference trigger is one of the triggers chosen to create the dataset being used and must be applied in the efficiency calculation to ensure all events have the same fired trigger. Otherwise, the dataset is comprised of a logical OR of various triggers and each event

may contain different fired triggers. Furthermore, when choosing a dataset to measure a particular trigger efficiency, that trigger must not be one of the triggers used to create that dataset. Therefore, an orthogonal dataset with respect to the trigger in question must be used.

In the present analysis, the trigger efficiency is observed to be different for events with no leptons and those with exactly one lepton [65]. For events with exactly one lepton, the trigger efficiency differs between events with one electron and events with one muon. Therefore, the trigger efficiency is measured separately for events with no leptons, those with exactly one electron, and those with exactly one muon.

For events with one muon, the SingleMu dataset is used, and the HLT_IsoMu24_eta2p1 trigger is used as the reference trigger. The HLT_IsoMu24_eta2p1 trigger is unprescaled and requires the leading muon candidate to have $p_{\mathrm{T}} > 24\,\mathrm{GeV}$ and $|\eta| < 2.1$. In addition to requiring exactly one muon, the offline analysis cuts applied on the SingleMu dataset are:

- at least three jets where the two leading jets have $p_{\mathrm{T}} > 70\,\mathrm{GeV}$;
- at least one primary vertex;
- angular cut $\Delta\hat{\phi}_{\mathrm{min}} > 4.0$;
- $E_{\mathrm{T}}^{\mathrm{miss}}$ cleaning.

The reason for applying these criteria is to ensure that the trigger efficiency is determined using events with the same topology as in the analysis.

The HLT_PFHT350_PFMET100 trigger efficiency for events with exactly one muon is shown by the red curve in Fig. 3.4. It is seen that the plateau region, or the region where the trigger becomes nearly fully efficient, occurs around 400 GeV. Although the HLT_PFHT350_PFMET100 trigger was sufficient for the 2011 analysis (see Ref. [65]), the analysis in 2012 (see Ref. [31]) is a so-called "shape" analysis, making use of the differences in shape between signal and background in the distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T, and the number of tagged b jets, and uses lower values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T to increase signal sensitivity. To recover efficiency in lower values of H_T, the unprescaled HLT_DiCentralPFJet50_PFMET80 analysis trigger was therefore investigated. The HLT_DiCentralPFJet50_PFMET80 trigger selects on at least two PF Jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$ and $|\eta| < 2.6$, PF $E_{\mathrm{T}}^{\mathrm{miss}} > 80\,\mathrm{GeV}$, and Calo $E_{\mathrm{T}}^{\mathrm{miss}} > 80\,\mathrm{GeV}$. While PF $E_{\mathrm{T}}^{\mathrm{miss}}$ at the trigger level is defined as in the HLT_PFHT350_PFMET100 trigger, Calo $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the negative vectorial transverse energy sum of Calo jets with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 3.0$. Unlike the HLT_PFHT350_PFMET100 trigger, the HLT_DiCentralPFJet50_PFMET80 has no requirement on H_T except for the implicit requirement of at least two PF jets. Using the SingleMu dataset and the unprescaled HLT_IsoMu24_eta2p1 trigger as the reference trigger, the HLT_DiCentralPFJet50_PFMET80 trigger is found to be fully efficient when the two leading jets each have $p_{\mathrm{T}} > 70\,\mathrm{GeV}$, as seen in Fig. 3.5. The offline analysis cuts for this efficiency curve include requiring PF $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$, at least one primary vertex, exactly one muon, and zero electrons.

Table 3.4: Data sets used to measure the trigger performance.

Dataset	Run range	$\int \mathcal{L}\, dt$ (pb^{-1})
/JetHT/Run2012A-13Jul2012-v1/AOD	190456-193621	807
/JetHT/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/JetHT/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/JetHT/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/JetHT/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/SingleMu/Run2012A-PromptReco-v1/AOD	190456-193621	800
/SingleMu/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/SingleMu/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/SingleMu/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/SingleMu/Run2012D-PromptReco-v1/AOD	203768-207469	7274
/SingleElectron/Run2012A-PromptReco-v1/AOD	190456-193621	800
/SingleElectron/Run2012B-13Jul2012-v1/AOD	193833-196531	4421
/SingleElectron/Run2012C-24Aug2012-v1/AOD	198022-198913	495
/SingleElectron/Run2012C-PromptReco-v2/AOD	198934-203002	6402
/SingleElectron/Run2012D-PromptReco-v1/AOD	203768-207469	7274
	Total	19399

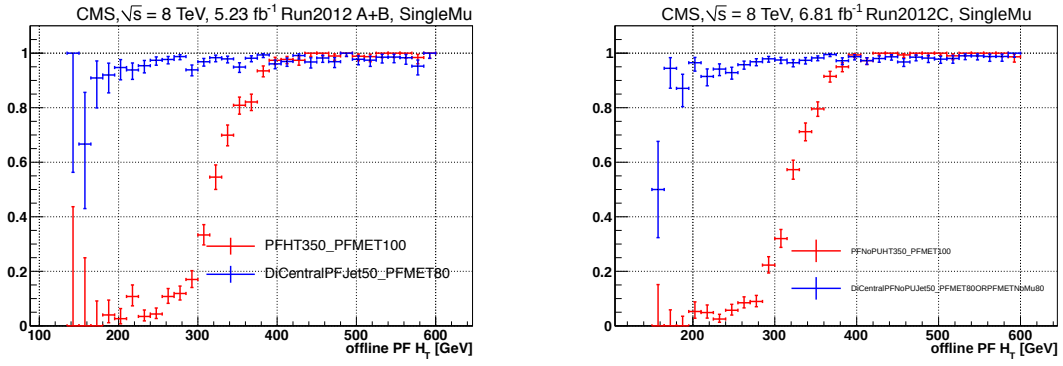


Figure 3.4: H_T trigger efficiency comparison for Run2012A,B (left) and Run2012C (right) between the HLT_PFHT350_PFMET100 only (red) and
HLT _Di Central PF Jet50 _PFMET80 only (blue) triggers. The efficiencies are measured using an orthogonal (single-muon) dataset. The offline selection requires at least two jets, PF $E_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$ to ensure PF $E_{\mathrm{T}}^{\mathrm{miss}}$ efficiency, and exactly one muon.

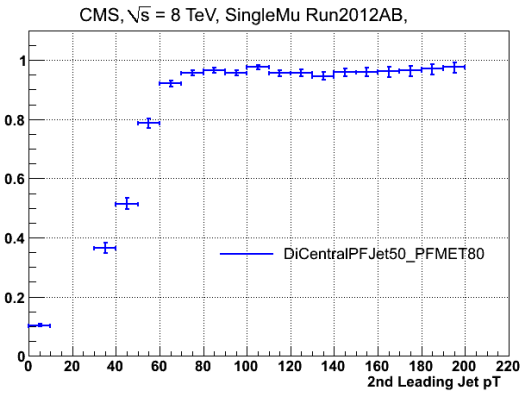


Figure 3.5: p_{T} distribution of second leading jet trigger efficiency in the SingleMu dataset. Offline selection requirements of PF $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, exactly one muon, and at least one primary vertex are applied.

The efficiencies of the HLT_PFHT350_PFMET100 and HLT_DiCentralPFJet50_PFMET80 triggers are compared in the SingleMu dataset, as seen in Fig. 3.4. There is virtually no turn-on region for the HLT_DiCentralPFJet50_PFMET80 trigger for $H_T > 150$ GeV, and this trigger is nearly fully efficient down to H_T values of 200 GeV. Throughout the entire H_T spectrum, this trigger comes close ($< 5\%$) to full efficiency, but never becomes fully efficient. The HLT_PFHT350_PFMET100 trigger is fully efficient for PF $H_T > 400$ GeV, so the HLT_PFHT350_PFMET100 and HLT_DiCentralPFJet50_PFMET80 triggers are combined in a logical OR to utilize the advantages of each trigger. Full efficiency is achieved for the HLT_PFHT350_PFMET100 trigger, and a higher trigger efficiency is obtained for $H_T < 400$ GeV with the HLT_DiCentralPFJet50_PFMET80 trigger.

A final analysis trigger, HLT_PFHT650, is considered. The trigger efficiency curve of this unprescaled trigger is shown in Fig. 3.6, measured in the SingleMu dataset. The trigger is fully efficient for tighter values of H_T, and requires no $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. Therefore, as long as

the H_T value is in the plateau region of this curve, there is no lower bound restriction on $E_{\mathrm{T}}^{\mathrm{miss}}$.

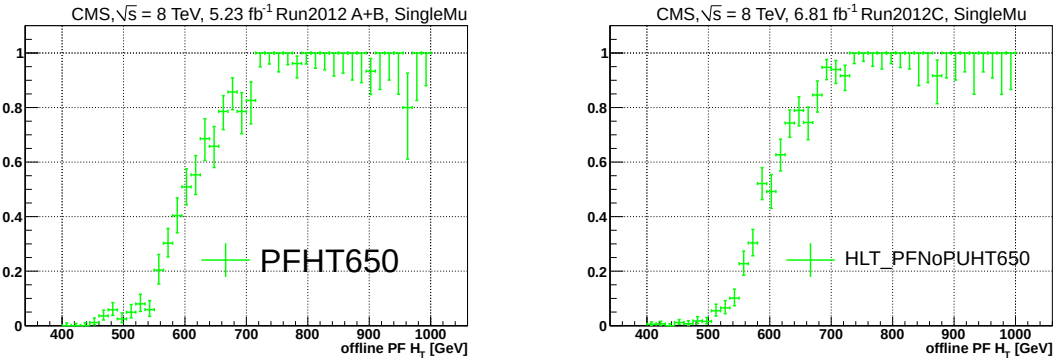


Figure 3.6: H_T trigger efficiency for Run2012A,B (left) and Run2012C (right) for the HLT_PFHT650 trigger. The efficiency is measured using an orthogonal (single-muon) dataset. The offline selection requires at least two jets, PF $E_{\mathrm{T}}^{\mathrm{miss}} > 200$ GeV, and exactly one muon.

The final trigger selection is thus a logical OR combination of HLT_PFHT350_PFMET100, HLT_DiCentralPFJet50_PFMET80, and HLT_PFHT650. These triggers exist in the Run2012A and Run2012B eras. Starting in the Run2012C era and continuing through the end of the Run2012D era, these triggers include pileup subtraction and changed names to HLT_PFNoPUHT350_PFMET100, HLT_DiCentralPFNoPUJet50_PFMETORPFMETNoMu80, and HLT_PFNoPUHT650, respectively. The run numbers and their respective run eras are given in Table 3.4. Additionally, the definition of the PF $E_{\mathrm{T}}^{\mathrm{miss}}$ leg of the HLT_DiCentralPFJet50_PFMET80 trigger was slightly modified to include the original PF $E_{\mathrm{T}}^{\mathrm{miss}}$ definition combined with a logical OR of PF $E_{\mathrm{T}}^{\mathrm{miss}}$ with muons removed. As can be seen in Fig. 3.7, the trigger efficiencies between

Run2012C and Run2012D are consistent and are thus combined. For the rest of this thesis, the Run2012A and Run2012B era trigger names will be used.

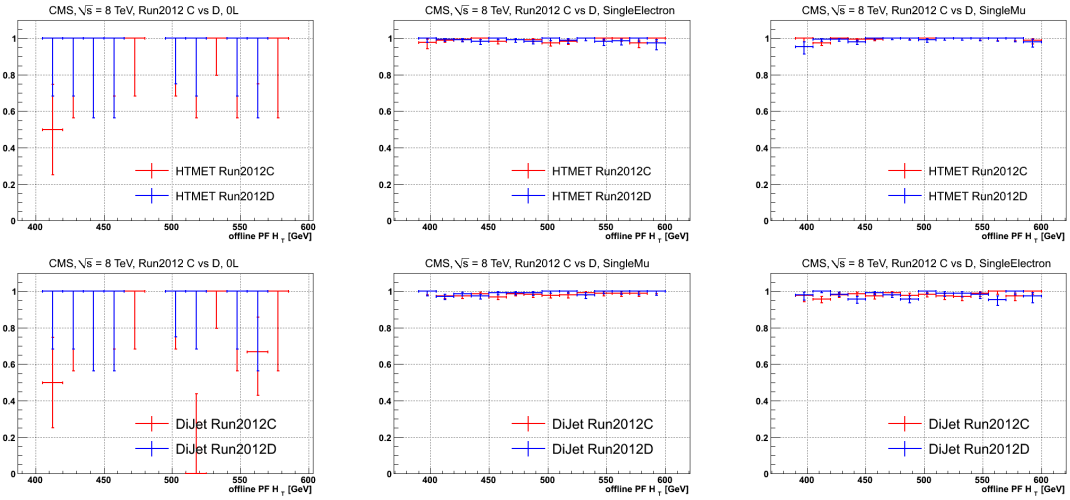


Figure 3.7: H_T trigger efficiency comparison between Run2012C (red) and Run2012D (blue), for the zero-lepton sample (left column), single-mu sample (middle column), and single-e sample (right column). The HLT_PFHT350_PFMET100 (top) and HLT_DiCentralPFJet50_PFMET80 (bottom) triggers are observed to be consistent between the two run eras.

For events with exactly one electron, the trigger efficiency is measured in the SingleElectron dataset. The offline analysis cuts are the same as those applied in the SingleMu dataset except exactly one electron and no muons are required. The unprescaled HLT_Ele27_WP80 trigger is used as a reference trigger. This trigger requires the leading electron to be isolated and have $p_{\mathrm{T}} > 27\,\mathrm{GeV}$.

For events with zero leptons, the JetHT dataset is used with the HLT_PHT350 trigger as the reference trigger. Because the HLT_PFHT350 has a relatively loose PF H_T threshold of $350\,\mathrm{GeV}$, it is prescaled, and statistical uncertainties are expected to be larger. This is confirmed from Fig. 3.8, which shows the final analysis trigger selection efficiencies for

events with zero leptons, one electron, and one muon. The offline analysis cuts for the zero lepton sample are the same as for the one electron and one muon samples except no leptons are required and $\Delta\hat{\phi}_{\min} < 4.0$ is imposed to increase the number of events.

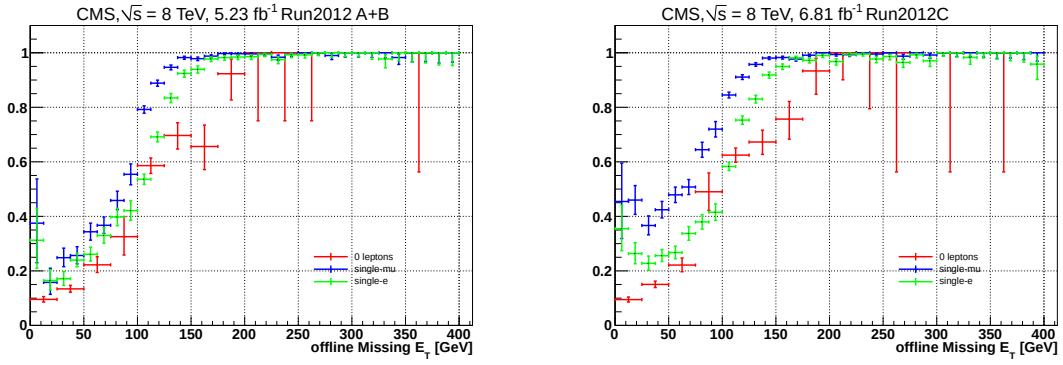


Figure 3.8: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency comparison between the QCD-dominated ZL sample (red), single-muon sample (blue), and single-electron sample (green). The plotted efficiency is for the OR of all three analysis triggers. The offline selection requires at least two jets and PF $H_T > 400\,\mathrm{GeV}$ to ensure PF H_T efficiency. Exactly one muon is required in the single-muon sample, and exactly one electron in the single-electron sample.

As can be seen in Fig. 3.9, the trigger efficiency is observed in simulation to differ between events producing genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ (such as $t\bar{t}$ and W+jets events) and events producing fake $E_{\mathrm{T}}^{\mathrm{miss}}$ (mostly QCD events). As seen in Fig. 3.10, the trigger efficiency for events with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ ($t\bar{t}$ and W+jets) are similar to $t\bar{t}$-enriched data events, by applying the $\Delta\hat{\phi}_{\min} < 4.0$ selection to the JetHT dataset. Similarly, the trigger efficiency for events with fake $E_{\mathrm{T}}^{\mathrm{miss}}$ is similar to that for the zero lepton (ZL) QCD-dominated sample, which is obtained by applying the $\Delta\hat{\phi}_{\min} > 4.0$ selection on data. Thus, the trigger efficiencies are calculated and applied separately for the ZL QCD- dominated sample and for the single-lepton (SL) sample (see Chapter 4 for the motivation behind choosing the ZL and

SL regions). For the SL sample, the weighted average of the one muon and one electron trigger efficiencies are computed.

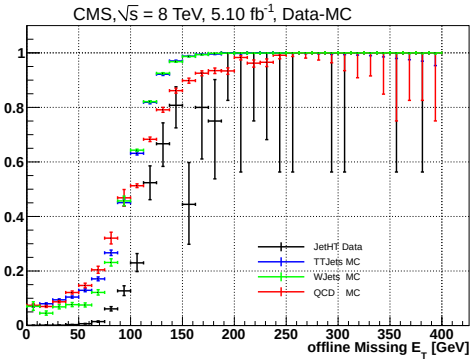


Figure 3.9: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in the ZL sample between MC events with genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ (blue and green) and fake $E_{\mathrm{T}}^{\mathrm{miss}}$ (red). This is compared with the $E_{\mathrm{T}}^{\mathrm{miss}}$ efficiency in data (black). Offline selection requirements of PF $H_T > 400$ GeV and PF $H_T < 500$ GeV are applied.

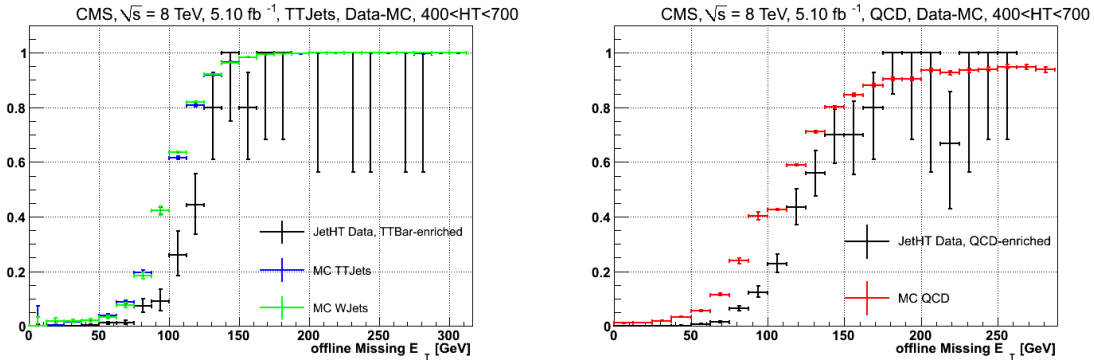


Figure 3.10: $E_{\mathrm{T}}^{\mathrm{miss}}$ analysis trigger efficiency comparison between MC and data. The data (black) is enriched with $t\bar{t}$ (blue) and W+jets (green) (left) or QCD (red, right) events. This data is compared to their respective MC processes. For data, the JetHT dataset is used. The offline selection requires $400\,\mathrm{GeV} < PFH_T < 700\,\mathrm{GeV}$, at least three PF jets, at least one primary vertex, and $PFE_{\mathrm{T}}^{\mathrm{miss}} > 200\,\mathrm{GeV}$.

Furthermore, studies with simulation show that the trigger efficiency is sensitive to the presence of at least one b-tagged jet, as seen in Fig. 3.11. The presence of a b-tagged jet is seen to have a greater impact on the zero-lepton sample. For the analysis bins not yet in the

plateau region, the requirement of at least one b-tagged jet is added to the offline analysis cuts when measuring the final trigger efficiency values.

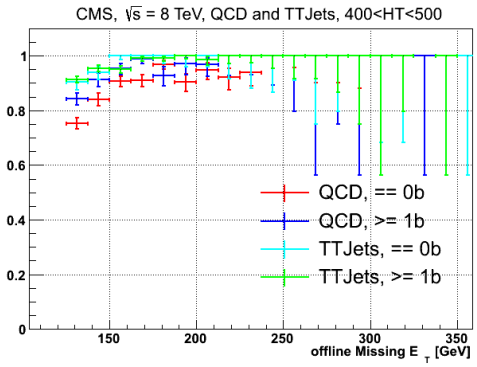


Figure 3.11: $E_{\mathrm{T}}^{\mathrm{miss}}$ trigger efficiency in simulation for QCD for events with no b-tagged jets (red) and events with at least one b-tagged jet (blue). Similar curves are shown for $t\bar{t}$ for events with no b-tagged jets (cyan) and events with at least one b-tagged jet (green).

In addition to splitting the trigger efficiency between SL and ZL QCD-dominated events, the trigger efficiencies are split into the 16 H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ bins used in the analysis (see Section 4). For bins already in the plateau region, where the trigger is fully efficient, bins are combined to reduce statistical uncertainty. Thus, the bins in the two highest H_T bins are combined, and the bins in the two highest $E_{\mathrm{T}}^{\mathrm{miss}}$ with two lowest H_T bins are combined.

The trigger efficiencies for the 16 bins for the zero-lepton QCD-dominated sample and the single-lepton sample are shown in Tables 3.5 and 3.6, respectively. The statistical uncertainties are included in the table.

The systematic uncertainties on the trigger efficiencies are discussed in the following section.

Table 3.5: Trigger efficiencies in the ZL QCD-dominated sample. The H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin ranges are in units of GeV. The uncertainties are statistical.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Selection (GeV)	$400 < H_T < 500$	$500 < H_T < 800$	$800 < H_T < 1000$	$H_T > 1000$
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	$0.86^{+0.07}_{-0.11}$	$0.67^{+0.09}_{-0.10}$	$1.00^{+0.00}_{-0.01}$	$1.00^{+0.00}_{-0.01}$
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	$0.88^{+0.08}_{-0.14}$	$1.00^{+0.00}_{-0.07}$	$1.00^{+0.00}_{-0.01}$	$1.00^{+0.00}_{-0.01}$
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	$1.00^{+0.00}_{-0.04}$	$1.00^{+0.00}_{-0.04}$	$1.00^{+0.00}_{-0.01}$	$1.00^{+0.00}_{-0.01}$
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	$1.00^{+0.00}_{-0.04}$	$1.00^{+0.00}_{-0.04}$	$1.00^{+0.00}_{-0.01}$	$1.00^{+0.00}_{-0.01}$

Table 3.6: Trigger efficiencies in the SL sample. The H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin ranges are in units of GeV. The uncertainties are statistical.

$E_{\mathrm{T}}^{\mathrm{miss}}$ Selection (GeV)	$400 < H_T < 500$	$500 < H_T < 800$	$800 < H_T < 1000$	$H_T > 1000$
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	$0.91^{+0.01}_{-0.01}$	$0.95^{+0.01}_{-0.01}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	$0.98^{+0.00}_{-0.01}$	$0.99^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$	$1.00^{+0.00}_{-0.00}$

3.7.2 Trigger Systematic Uncertainties

The trigger systematic uncertainties are due to the difference in event composition between the region where the trigger efficiencies are measured, and the region where the trigger efficiencies are applied. In this section, we study the effect of such differences on the $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T efficiencies for the case of the zero-lepton QCD-dominated and single-lepton samples in each of the sixteen $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T bins.

The sample composition of each bin is split into two components: the QCD fraction and non-QCD fraction. Table 3.7 shows the fraction of MC QCD in the regions where the efficiencies are applied. The non-QCD fraction for these regions is the difference between the QCD fraction and unity. Table 3.8 shows the QCD fraction of the regions where the efficiencies are used.

Table 3.7: QCD fraction in the LDP and SL samples with at least 1 b jet requirement. The H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin ranges are in units of GeV.

	$400 < H_T < 500$	$500 < H_T < 800$	$800 < H_T < 1000$	$H_T > 1000$
LDP				
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.92	0.94	0.94	0.96
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.82	0.86	0.87	0.90
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.42	0.64	0.75	0.78
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.00	0.29	0.41	0.58
SL				
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.00	0.00	0.02	0.00
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.00	0.01	0.00	0.00
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.00	0.00	0.00	0.00
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.00	0.00	0.00	0.00

These efficiencies for QCD and non-QCD processes measured in data are shown in Tables 3.5 and 3.6, respectively. Using these efficiencies, the "process-weighted" $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T efficiencies are computed in each of the regions in Table 3.8:

$$\epsilon = f_{QCD} \cdot \epsilon_{QCD} + f_{other} \cdot \epsilon_{other}, \tag{3.5}$$

where f_{QCD} and f_{other} are the corresponding fractional contributions of QCD and non-QCD, respectively, and ϵ_{QCD} and ϵ_{other} are the efficiencies in Table 3.5 and 3.6, respectively. The difference in the "process-weighted" efficiency between the signal region and the trigger-efficiency-measured region is used as the systematic uncertainty. These systematic uncertainty results are shown in Table 3.9. In the bins where ϵ_{QCD} and ϵ_{other} are equal, this procedure returns a systematic uncertainty of zero. In this case, an uncertainty of 1% is assigned to cover residual effects.

Table 3.8: QCD fractional composition in $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T bins of the LDP and SL samples where the trigger efficiencies are used. At least one b jet is required. The H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin ranges are in units of GeV.

LDP	$400 < H_T < 500$	$500 < H_T < 800$	$800 < H_T < 1000$	$H_T > 1000$
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.93	0.93	0.94	0.96
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.86	0.88	0.85	0.86
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.64	0.62	0.72	0.75
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.00	0.35	0.46	0.57
SL				
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.00	0.00	0.02	0.00
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.00	0.00	0.00	0.00
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.00	0.01	0.00	0.00
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.00	0.00	0.00	0.00

Table 3.9: Trigger systematics in the LDP and SL samples. Where no systematic is found, a minimum uncertainity of 0.01 is assigned to cover residual effects. The H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin ranges are in units of GeV.

LDP	$400 < H_T < 500$	$500 < H_T < 800$	$800 < H_T < 1000$	$H_T > 1000$
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.01	0.01	0.01	0.01
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.01	0.01	0.01	0.01
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.01	0.01	0.01	0.01
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.01	0.01	0.01	0.01
SL				
$125 < E_{\mathrm{T}}^{\mathrm{miss}} < 150$	0.01	0.01	0.01	0.01
$150 < E_{\mathrm{T}}^{\mathrm{miss}} < 250$	0.01	0.01	0.01	0.01
$250 < E_{\mathrm{T}}^{\mathrm{miss}} < 350$	0.01	0.01	0.01	0.01
$E_{\mathrm{T}}^{\mathrm{miss}} > 350$	0.01	0.01	0.01	0.01

Compared to Tables 3.5 and 3.6, it is seen that the dominant uncertainties on the trigger efficiencies are the statistical uncertainties.

Chapter 4

Background Evaluation Methods

After applying the baseline selection criteria discussed in Section 3.5.4, which is referred to as the zero-lepton (ZL) selection, the main SM background is from the combined category with a top-antitop quark pair ($t\bar{t}$), a W boson produced in association with jets ($W+\mathrm{jets}$), or a single-top quark. This category is referred to as "ttWj." These backgrounds produce genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ because the W boson from these events can decay into a neutrino and a charged lepton (top quarks decay to a W boson in association with a b jet). The events from these processes are considered in one category, since they are characterized by typically having at least one lepton and genuine $E_{\mathrm{T}}^{\mathrm{miss}}$. Choosing a zero-lepton sample already greatly reduces this background, but background events from this category remain due to electrons or muons that are not identified, not isolated, or that lie outside the fiducial acceptance of the analysis, as well as to W bosons that decay to a tau lepton that decays hadronically. Most events in this background category that have electrons or muons are selected because

the lepton p_{T} value lies below the p_{T} threshold of 10 GeV required for electron or muon identification.

Multijet quantum chromodynamic (QCD) events, produced via strong interactions, are also a major background due to the mismeasurement of jet transverse moment p_{T} resulting in significant "fake" $E_{\mathrm{T}}^{\mathrm{miss}}$. QCD can also produce genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ when the charm or bottom quarks decay semileptonically, but genuine $E_{\mathrm{T}}^{\mathrm{miss}}$ in these events is typically smaller than fake $E_{\mathrm{T}}^{\mathrm{miss}}$ produced through jet p_{T} mismeasurement. In a QCD event, the two largest-p_{T} jets are often nearly back-to-back. When one of these jets is mismeasured, the mismeasured jet p_{T} is reflected in the computation of $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, one characterization of QCD events is that $E_{\mathrm{T}}^{\mathrm{miss}}$ tends to align with the mismeasured jet. The $\Delta\hat{\phi}_{\mathrm{min}}$ variable, which is related to the angle between jets and $E_{\mathrm{T}}^{\mathrm{miss}}$, provides a powerful discriminant between QCD events and non-QCD events and is only weakly correlated with $E_{\mathrm{T}}^{\mathrm{miss}}$, as discussed in Section 3.5.1.

The final significant background to this analysis is from events with a Z boson produced in association with jets (Z+jets). The Z boson can decay to two neutrinos ($\mathrm{Z} \rightarrow \nu\bar{\nu}$) and produce genuine $E_{\mathrm{T}}^{\mathrm{miss}}$. $\mathrm{Z} \rightarrow \nu\bar{\nu}$ events are "irreducible" because their signature is the same as signal events, i.e. they both have multiple jets, b jets, and $E_{\mathrm{T}}^{\mathrm{miss}}$.

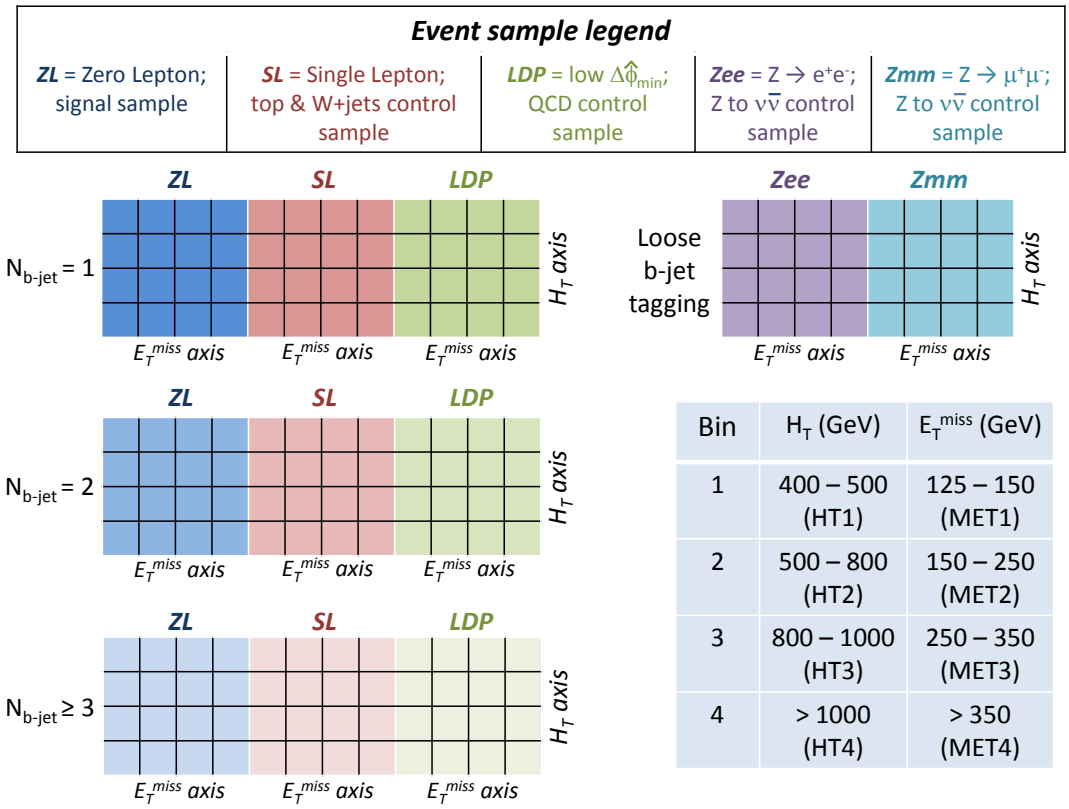
Data-driven background estimation techniques are used to limit reliance on the MC. If the MC does not model the data well, strong reliance on MC can give unrealistic results. However, MC can be used to validate the background estimation techniques and provide "closure" tests. A closure test compares the background estimate derived by applying the

estimation method to simulated events, treated like data, to the "true" background event numbers known from the simulation. Closure tests are used to validate the estimation procedures and to assign associated systematic uncertainties

A background "control" sample is used to mimic the background content of the ZL signal sample, and is dominated by the background in question using a set of selection cuts that are orthogonal to the nominal selection cuts. The orthogonality allows the nominal analysis cuts to be unbiased. Validation tests in MC are performed to show closure between the control sample and MC signal sample.

This chapter discusses techniques to estimate the major SM backgrounds, starting with the $\mathrm{Z} \rightarrow \nu\bar{\nu}$ background. This $\mathrm{Z} \rightarrow \nu\bar{\nu}$ background estimation section also includes a detailed discussion investigating the use of a photon control sample. The idea of this sample was to decrease statistical uncertainties, but it ultimately was not used because of data-MC differences. A summary of the top and QCD background estimation techniques is discussed in the section thereafter. Other backgrounds, such as those from diboson and Drell-Yan processes, contribute a negligible amount. Except for the description of the photon control sample, the SM background estimation techniques are documented in detail in Ref. [66] and are only briefly summarized in this chapter.

The final binning is shown as a schematic diagram in Fig. 4.1. For the ttWj and QCD backgrounds, the final binning is chosen to be four bins in H_T, four bins in $E_{\mathrm{T}}^{\mathrm{miss}}$, and three bins in b-jet multiplicity for the ZL signal sample and background control sample. For the $\mathrm{Z} \rightarrow \nu\bar{\nu}$ background, the same four bins in H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ are used, but without the



Bin	H_T (GeV)	E_T^{miss} (GeV)
1	400 – 500 (HT1)	125 – 150 (MET1)
2	500 – 800 (HT2)	150 – 250 (MET2)
3	800 – 1000 (HT3)	250 – 350 (MET3)
4	> 1000 (HT4)	> 350 (MET4)

Figure 4.1: Schematic diagram illustrating the 176 mutually exclusive observables of the analysis. There are four bins each of $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_T. The three bins for $N_{\mathrm{bjet}} = 1$, 2, and ≥ 3 are used for the top and QCD backgrounds. The $\mathrm{Z} \to \nu\bar{\nu}$ background uses a $\mathrm{Z} \to \ell^+\ell^-$ control sample without the additional b-jet-multiplicity bins. The designations HTi and METi ($i = 1 - 4$) are used to label the individual H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bin divisions. From Ref. [31].

additional b jet multiplicity bins. The control sample for the $\mathrm{Z} \to \nu\bar{\nu}$ background is formed by selecting $\mathrm{Z} \to \ell^+\ell^-$ events, where (ll) are either both electrons or both muons. This is described in Section 4.1. The single-lepton (SL) sample is used as the control sample for $t\bar{t}$, W+jets, and single-top events, which is described in Section 4.2. The low-delta-phi (LDP) sample, $\Delta\phi_N < 4$, is used as a control sample for the QCD background, and is discussed in Section 4.2.2.

The event counts, or observables, of the 176 mutually exclusive bins in Fig. 4.1 are simultaneously analyzed in a global likelihood fit. The likelihood function is described in Section 5, but the formulation of the control and signal sample for each background, and how they fit into the likelihood, are now briefly discussed. In particular, each background component contributes observables, or Poisson mean parameters, used in the global likelihood fit. These observables will be derived for each background discussed below.

4.1 $\mathrm{Z} \to \nu\bar{\nu}$ Background

The $\mathrm{Z} \to \nu\bar{\nu}$ control sample is created by selecting events where the Z boson decays to two leptons ($\mathrm{Z} \to \ell^+\ell^-$), where the two leptons (ll) are either both electrons or both muons. The $\mathrm{Z} \to \ell^+\ell^-$ control sample, with the two leptons "ignored," mimic the event kinematics in $\mathrm{Z} \to \nu\bar{\nu}$ events. The two leptons are ignored by recomputing $E_{\mathrm{T}}^{\mathrm{miss}}$ and other physics objects in the event as if the two leptons do not exist. The $\mathrm{Z} \to \ell^+\ell^-$ control sample is formed by using the DoubleElectron and DoubleMu datasets, which use triggers requiring either two electrons or muons, respectively, and then selecting events where the dileptons form an invariant mass close (within 15 GeV) to the Z boson (91.2 GeV). The event yields for $\mathrm{Z} \to \mathrm{e}^+\mathrm{e}^-$ and $\mathrm{Z} \to \mu^+\mu^-$ are computed for each of the 16 bins shown in Fig. 4.1, which are corrected for background, acceptance (measured in MC), and detection efficiency (derived in data).

Since the $\mathrm{Z} \rightarrow \ell^+\ell^-$ event yields are small, the event rate is increased by loosening the requirements on the b jet. Thus, the b-jet CSV value is relaxed from 0.679 to 0.244. An extrapolation factor, $\mathcal{F}_{Z\nu\nu}$, is used to scale the observed events in the control sample to measure the event yields in the signal sample. The extrapolation factor is the probability that an event with a loose b-tag requirement also passes the nominal b-tagging selection, and is defined as the ratio of the sum of the number of $\mathrm{Z} \rightarrow \ell^+\ell^-$ events over all H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bins that satisfy the nominal b-jet tagging requirements to the number of $\mathrm{Z} \rightarrow \ell^+\ell^-$ events in all H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ bins that contain at least one jet with loose b-jet tagging requirements. Each bin in Fig. 4.1 is scaled by the extrapolation factor, and there is one extrapolation factor for each b-jet bin.

The higher the CSV selector value, the more likely is a jet to be a b jet. The CSV selector for the most b-like jet, the second most b-like jet, and the third most b-like jet is shown in Fig. 4.2. The blue curves show the results for data events selected with the ZL sample requirements except with the requirement on the $\Delta\hat{\phi}_{\mathrm{min}}$ variable reversed, $\Delta\hat{\phi}_{\mathrm{min}} < 4.0$, and without a b-tagging requirement. For b-tagging discriminant values above the loose-tagging threshold of 0.244, the distributions are seen to be similar to those of $\mathrm{Z} \rightarrow \ell^+\ell^-$ events (points with error bars), and to have far better statistical precision. Therefore, the $\Delta\hat{\phi}_{\mathrm{min}} < 4.0$ sample is used as a control sample to verify the H_T and $E_{\mathrm{T}}^{\mathrm{miss}}$ dependence of the b-tagging extrapolation factors. The red curves in Fig. 4.2 show the b-tagging discriminant values for events selected with the signal-sample (ZL) requirements, except without the b-tagging.

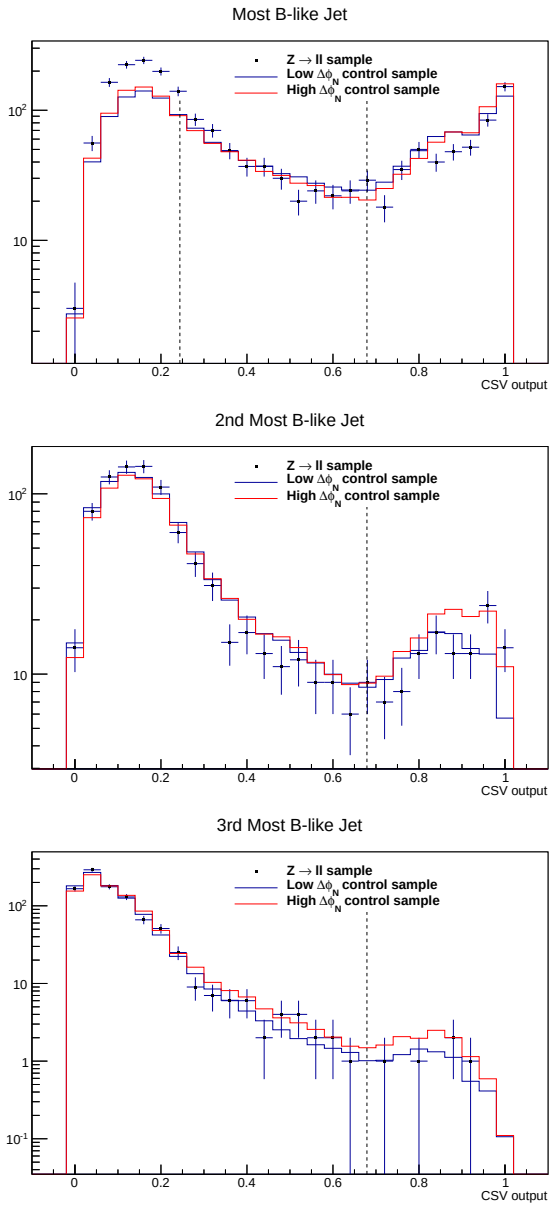


Figure 4.2: CSV distributions for the most b-like (top), second most b-like (center) and third most b-like (bottom) jet for Z $\rightarrow \ell^+\ell^-$ data events (black points), data events selected with the ZL sample requirements except with $\Delta\hat{\phi}_{\rm min} < 4.0$ and without a b-tagging requirement (blue), and Z $\rightarrow \ell^+\ell^-$ data events with $\Delta\phi_N < 4$ (blue), and events selected with the signal-sample (ZL) requirements, except without the b-tagging (red). The CSV shapes between all samples are similar. The dashed lines on the left plot indicate the loose b-jet tagging (left-dashed line) and nominal b-tagging (right-dashed line) requirements. For the CSV distributions of the second and third most b-like jet, the dashed lines indicate the nominal b-tagging requirement. From Ref. [66].

The extrapolation factor $\mathcal{F}_{Z\nu\nu}$ is found to be independent of H_T and $E_{\rm T}^{\rm miss}$ for $N_{b\rm jet} = 2$ and $N_{b\rm jet} \geq 3$. A variation in $E_{\rm T}^{\rm miss}$ up to 25% is found for $N_{b\rm jet} = 1$ events and is accounted for in the $N_{b\rm jet} = 1$ extrapolation factors.

Let the indices i and j represent $E_{\rm T}^{\rm miss}$ and H_T bins, respectively. The estimated number of Z+jets events in the ZL sample is given by

$$\mu^{zll}_{zll;\,i,j} \;=\; \left(\mu^{Z\nu\nu}_{ZL;\,i,j,1} \cdot S_{ll} \cdot A_{ll;\,i} \cdot \epsilon_{ll}\right) / \left(\mathcal{F}_{Z\nu\nu;\,1} \cdot R_B\right) \tag{4.1}$$

where $\ell =$ e or μ, $\mu^{Z\nu\nu}_{ZL;\,i,j,1}$ is the $\mathrm{Z} \to \ell^+\ell^-$ yield in the $i-j$ bin, $A_{\ell\ell;\,i}$ are the acceptances for the $\mathrm{Z} \to \ell^+\ell^-$ samples, $\epsilon_{\ell\ell}$ are the efficiencies for the $\mathrm{Z} \to \ell^+\ell^-$ samples, $S_{\ell\ell}$ is a scale factor accounting for the systematic uncertainty scale factors, $R_B = 5.95 \pm 0.02$ is the ratio of the $\mathrm{Z} \to \nu\bar{\nu}$ to the $\mathrm{Z} \to \ell^+\ell^-$ branching fraction [10], and $\mathcal{F}_{Z\nu\nu;\,1}$ is the extrapolation factor relating the $N_{b\rm jet} = 1$ selection efficiency to the efficiency of the loosened b-jet tagging requirement. A more detailed derivation of these variables can be found in Ref. [66].

Let k represent the $N_{b\rm jet}$ bin index. The estimated number of the Z+jets events in the ZL sample for $N_{b\rm jet} = 2$ and $N_{b\rm jet} \geq 3$ ($\mu^{Z\nu\nu}_{ZL;\,i,j,k}$) are defined by the $N_{b\rm jet} = 1$ result multiplied by the ratio of b-jet tagging extrapolation factors:

$$\mu^{Z\nu\nu}_{ZL;\,i,j,k} \;=\; \mu^{Z\nu\nu}_{ZL;\,i,j,1} \cdot \left(\mathcal{F}_{Z\nu\nu;\,k} / \mathcal{F}_{Z\nu\nu;\,1}\right) \;, \tag{4.2}$$

The Poisson mean parameters, or observables for this background, are given by

$$n_{zll;\,i,j,k} \;=\; \mu^{zll}_{ZL;\,i,j,k}/P_{ll} \tag{4.3}$$

where $n_{zll;\,i,j,k}$ is the number of Z $\rightarrow \ell^+\ell^-$ events in the i-j-k bin, and P_{ll} is the purity $(S/(S+B))$ of the Z $\rightarrow \ell^+\ell^-$ sample.

4.1.1 Photon + Jets Control Sample

Though using Z $\rightarrow \ell^+\ell^-$ events as a control sample is a clear and direct way to estimate the Z $\rightarrow \nu\bar{\nu}$ background, an alternative method to estimate the Z $\rightarrow \nu\bar{\nu}$ background, explored by a number of SUSY analyses within CMS (see Refs. [50, 18, 36]), is to use photons + jets (γ+jets) events as a control sample. At high vector boson p_{T} much larger than the mass of the Z (p_{T} above 150 GeV), event kinematics between Z $\rightarrow \nu\bar{\nu}$ and γ+jets are almost the same if the photon in the γ+jets events are ignored [3]. Due to boson-quark couplings, the cross sections between the two processes are different and the γ+jets events have a larger production cross section. As discussed in Section 4.1, the Z $\rightarrow \ell^+\ell^-$ control sample event yields are increased by loosening the b-jet tagging requirements. By using a photon control sample, these yields might be increased even further, providing a complementary technique to evaluate the Z $\rightarrow \nu\bar{\nu}$ background with good statistical precision. Therefore, we investigated using the γ+jets event technique to evaluate the Z $\rightarrow \nu\bar{\nu}$ background for this analysis. We were asked to investigate this method by the SUSY conveners and other

reviewers. Although the γ+jets event technique was ultimately not used in the analysis for the reason discussed below, we nonetheless document the studies performed here. The main work performed for this thesis concerned the selection of the γ+jets event sample and in validating the selection.

The SinglePhoton2012C dataset was used for the data photon sample, which corresponded to an integrated luminosity of 6.4 fb^{-1}. This dataset includes a logical OR combination of single-photon triggers, each having a different photon-p_{T} requirement. The HLT_Photon135 trigger is used, which requires at least one isolated photon whose p_{T} is greater than 135 GeV. To ensure full trigger efficiency with respect to the baseline selection, a photon candidate is required to have p_{T} of at least 140 GeV. Photons are selected based on the critieria recommended by the SUSY Photon Group, which is outlined in Ref. [29], and summarized in Table 4.1. The selection criteria values differ depending on whether the photon is found in the barrel region or the endcaps.

Table 4.1: Photon selection criteria. All values are upper bounds for the corresponding variable. $pho_{p_{\mathrm{T}}}$ is the p_{T} of the photon candidate.

Variable	Barrel (Endcaps)
Conversion safe electron veto	Yes (Yes)
Single tower H/E	0.05 (0.05)
$\sigma_{i\eta i\eta}$	0.012 (0.034)
Rho corrected PF charged hadron isolation	2.6 (2.3)
Rho corrected PF neutral hadron isolation	3.5 + 0.04$*pho_{p_{\mathrm{T}}}$ (2.9 + 0.04$*pho_{p_{\mathrm{T}}}$)
Rho corrected PF photon isolation	1.3 + 0.0005$*pho_{p_{\mathrm{T}}}$ (none)

Photon and electrons look similar in the calorimeter, so a "conversion-safe electron veto" distinguishes electrons and photons by checking if the reconstructed photon candi-

date can be mapped back to clusters of hits in the ECAL, and is therefore an electron. The "conversion-safe" portion refers to checking the presence of matched "conversion" electrons, (see Ref. [30]), which occur when a photon converts into an electron-positron pair. If an electron is found by mapping to the ECAL, but is a conversion electron, the photon candidate is not vetoed. Thus, a conversion-safe electron veto is applied to the reconstructed photon candidate to ensure that it is not actually an electron that is a conversion electron. H/E is the energy of an HCAL tower within a cone of Δ_R = 0.15 around the reconstructed photon, divided by the energy due to the cluster of hits in the ECAL. Since most of the photon shower will be in the ECAL, the H/E value of the photon candidate should be low. This variable also prevents selecting on jets that otherwise appear similar to photons. A cut on the shower shape variable, $\sigma_{i\eta i\eta}$ [23] is required, and is related to the η width of the cluster of hits. This width is typically larger for photons from π^0 decay, if the two photons in the $\pi^0 \rightarrow \gamma\gamma$ decay become merged, than for isolated photons. The charged hadron isolation is the p_{T} sum of all charged hadrons that have tracks coming from the primary vertex within Δ_R = 0.3 around the reconstructed photon candidate. The rho correction refers to the correction for the presence of additional proton-proton collisions using the event energy density ($\frac{dE}{d\eta}$). The neutral hadron and photon isolation variables are similar to the charged isolation variable, except that neutral hadrons and photons are considered instead of charged hadrons.

The photon control sample is selected by requiring at least one photon in the event, and then re-calculating $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T, jet multiplicity, and $\Delta\hat{\phi}_{\mathrm{min}}$ with the highest-p_{T} photon

removed. To recompute $E_{\mathrm{T}}^{\mathrm{miss}}$ with the photon removed, the photon momentum can simply be added to the original $E_{\mathrm{T}}^{\mathrm{miss}}$ value. For H_T, jet multiplicity, and $\Delta\hat{\phi}_{\mathrm{min}}$, the jet collection must be cleaned of photons since the PF jet collection includes photon candidates that are treated as jets. Therefore, a jet is removed from the PF jet collection if a jet is found within Δ_R = 0.20 around a reconstructed photon candidate.

The number of event counts for the photon control sample, using the SinglePhoton2012C dataset, is shown in Table 4.2 for the 16 analysis bins. The values shown include exactly one nominal (loose) b-tagged jet, and are after applying the baseline selection.

Table 4.2: Number of event counts for the photon control sample in the SinglePhoton2012C dataset after requiring exactly one nominal (loose) b-tagged jet and applying the baseline selection.

	400$<H_T<$500	500$<H_T<$800	800$<H_T<$1000	$H_T>$1000
Exactly 1 b-jet				
125$<E_{\mathrm{T}}^{\mathrm{miss}}<$150	99(238)	84(191)	8(12)	4(10)
150$<E_{\mathrm{T}}^{\mathrm{miss}}<$250	544(1276)	449(1021)	46(86)	26(50)
250$<E_{\mathrm{T}}^{\mathrm{miss}}<$350	157(389)	152(342)	24(49)	11(30)
$E_{\mathrm{T}}^{\mathrm{miss}}>$350	78(36)	92(216)	23(37)	13(28)

The CSV selector distributions for the γ+jets control sample for data, MC events, and $\mathrm{Z} \to \nu\bar{\nu}$ MC events are shown in Fig. 4.3. It is seen that the shapes of all curves are similar. However, when compared with Fig. 4.2, the shapes are not similar. It is seen that the CSV selector distribution shape for $\mathrm{Z} \to \ell^+\ell^-$ data events are not similar to the CSV selector distribution shape for $\mathrm{Z} \to \nu\bar{\nu}$ MC events, γ+jets MC events with the leading photon removed, or γ+jets data events with the leading photon removed. Thus, even though the CSV output variable has a similar shape for the γ+jets sample in data and MC, and even

though the MC shape for Z $\rightarrow \nu\bar{\nu}$ events is also similar, the MC shape for Z $\rightarrow \nu\bar{\nu}$ events is not the same as the data shape for Z $\rightarrow \ell^+\ell^-$ events, as can be seen in Fig. 4.4. However, it is well known that the theory for Z+jets events where one or more of the jets is a b jet is incomplete. Previous studies using the γ+jets method to determine the Z $\rightarrow \nu\bar{\nu}$ background did not require b jets. We interpret the lack of agreement between the CSV discriminator output variable between simulated Z $\rightarrow \nu\bar{\nu}$ events and data Z $\rightarrow \ell^+\ell^-$ events to be due to the theoretical incompleteness of the Z+jets simulation when b jets are required. In the end, we demonstrated to the satisfaction of the SUSY group that the γ+jets method was not appropriate as a means to determine the Z $\rightarrow \nu\bar{\nu}$ analysis in our study. This shows the importance of using data-driven background methods, since estimates based on MC can be unreliable.

4.2 $t\bar{t}$, W+jets, Single Top, and QCD background

The remaining main SM background estimation techniques are now discussed. A detailed description can be found in Refs. [66, 31], but a brief summary is given here. The formulations of the control samples are discussed, and the expressions for the Poisson parameters are given.

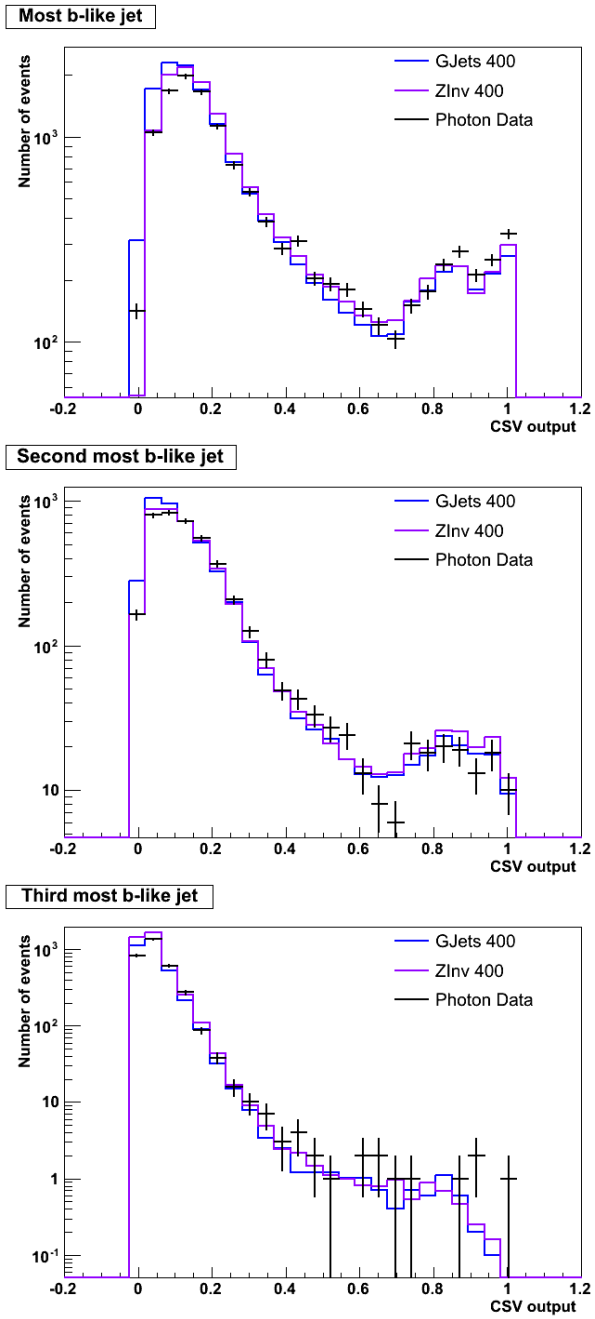


Figure 4.3: CSV distributions for the most b-like (top), second most b-like (center) and third most b-like (bottom) jet for γ+jets data events (black points), γ+jets simulated events (blue), and $Z \rightarrow \nu\bar{\nu}$ simulated events (red). The CSV shapes between all samples are similar. The dashed lines on the left plot indicate the loose b-jet tagging (left-dashed line) and nominal b-tagging (right-dashed line) requirements. For the CSV distributions of the s econd and third most b-like jet, the dashed lines indicate the nominal b-tagging requirement. γ+jets simulated events are normalized to the integrated luminosity given by γ+jets data events, and simulated $Z \rightarrow \nu\bar{\nu}$ events are normalized to an arbitrary integrated luminosity.

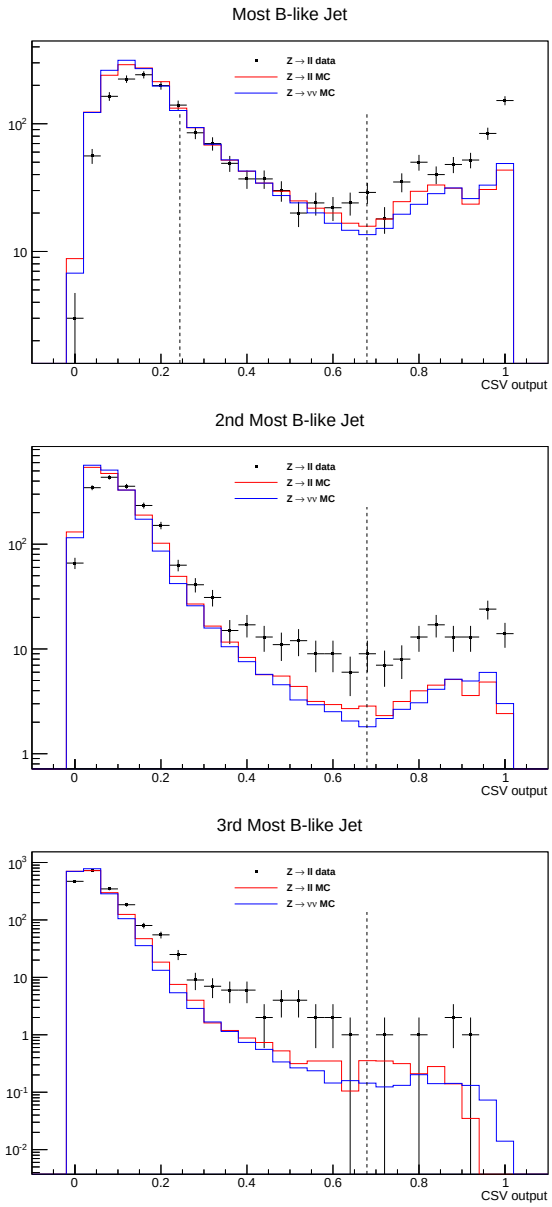


Figure 4.4: CSV distributions for the most b-like (top), second most b-like (center) and third most b-like (bottom) jet for $Z \to \ell^+\ell^-$ data events (black points), $Z \to \ell^+\ell^-$ simulated events (red), and $Z \to \nu\bar{\nu}$ simulated events (blue). The CSV shapes between the $Z \to \ell^+\ell^-$ and $Z \to \nu\bar{\nu}$ simulated events are similar, but they are not similar to the $Z \to \ell^+\ell^-$ data events. The dashed lines on the left plot indicate the loose b-jet tagging (left-dashed line) and nominal b-tagging (right-dashed line) requirements. For the CSV distributions of the second and third most b-like jet, the dashed lines indicate the nominal b-tagging requirement. These series of plots were made by Troy Mullohand, a member of the RA2b group, in private collaboration.

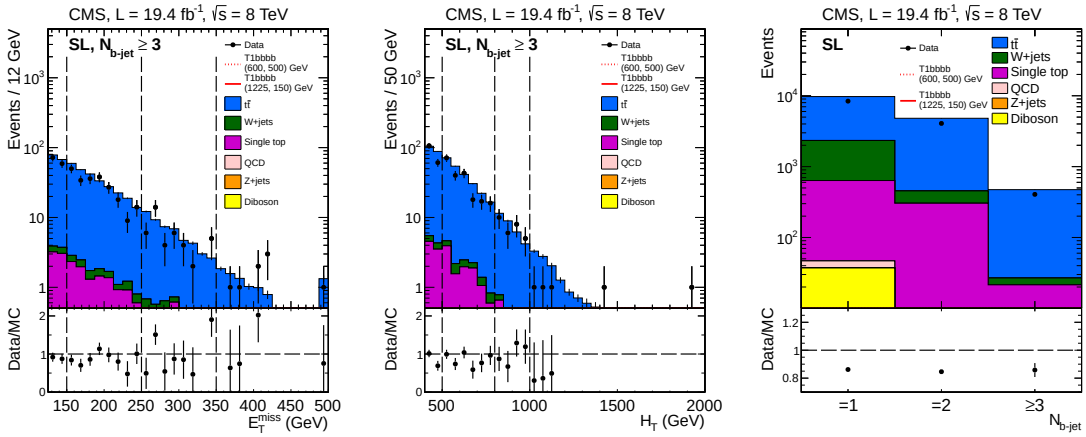


Figure 4.5: Data and MC distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (left), H_T (middle), and $N_{b\mathrm{jet}}$ (right) for the SL control sample. The accompanying ratio plots at the bottom show the ratio of data to simulated standard model events. Dashed vertical lines indicate the $E_{\mathrm{T}}^{\mathrm{miss}}$ or H_T bin divisions. From Ref. [31].

4.2.1 $t\bar{t}$, W+jets, and single-top backgrounds

The $t\bar{t}$, single-top, and W+jets background (ttWj) events are characterized by having a W boson that decays to a lepton and a neutrino. The neutrino provides genuine $E_{\mathrm{T}}^{\mathrm{miss}}$. Thus, requiring exactly one electron or muon in the event gives a sample that is dominated by $t\bar{t}$ events, with small contribution from W+jets and single-top events. This single-lepton (SL) sample is illustrated in Fig. 4.5 and is used as a control sample for the mixture of $t\bar{t}$, W+jets, and single-top backgrounds. Contributions from Z+jets, QCD, and diboson events are at the level of around 1% and are accounted for through a systematic uncertainty. The distributions of two representative T1bbbb scenarios, one with $(m_{\mathrm{gluino}}, m_{\mathrm{LSP}}) = (600\,\mathrm{GeV}, 500\,\mathrm{GeV})$ and the other with $(m_{\mathrm{gluino}}, m_{\mathrm{LSP}}) = (1225\,\mathrm{GeV}, 150\,\mathrm{GeV})$, are also included. These $(m_{\mathrm{gluino}}, m_{\mathrm{LSP}})$ points are at the limit of our expected sensitivity (see Chapter 5).

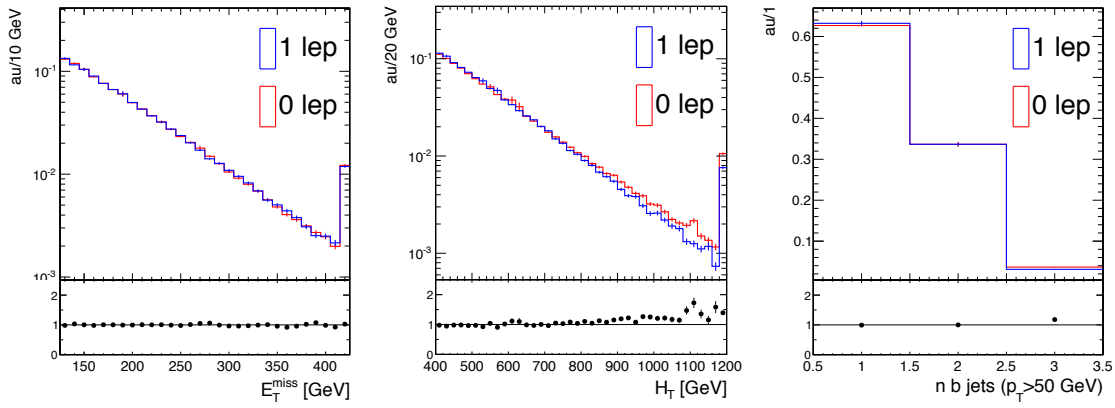


Figure 4.6: $E_{\mathrm{T}}^{\mathrm{miss}}$ (left), H_T (middle), and $N_{b\mathrm{jet}}$ (right) distribution comparing SL (blue) to ZL (red) selection for $t\bar{t}$, W+jets, and single-top events. $H_T \geq 400$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}} \geq 125$ GeV selections have been applied. From [31].

Extensive verification of the compatibility of the SL sample with the ZL sample for $t\bar{t}$, W+jets, and single-top events is presented in a previous CMS Note [65], where the robustness of the method is demonstrated.

As seen in Figure 4.6, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for the SL MC samples are compatible with the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions from the nominal zero-lepton (ZL) MC samples. This validates use of the SL control sample to model the ZL sample for $t\bar{t}$, W+jets, and single-top background. Thus, the SL control sample describes the shape of $t\bar{t}$, W+jets, and single-top background in the three dimensions of $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T, and $N_{b\mathrm{jet}}$, and provides a three-dimensional (3D) histogram-probability-density function (PDF).

The top row of Fig. 4.7 show the ratio of ZL-to-SL ttWj events in MC for each H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. Each plot represents a different $N_{b\mathrm{jet}}$ bin, and the distributions are for $N_{b\mathrm{jet}} = 1$, 2, and ≥ 3 from left to right, respectively. Thus, there are 16 points shown for each plot, and there are three plots. The bottom row shows the ratios in the top plot divided by their

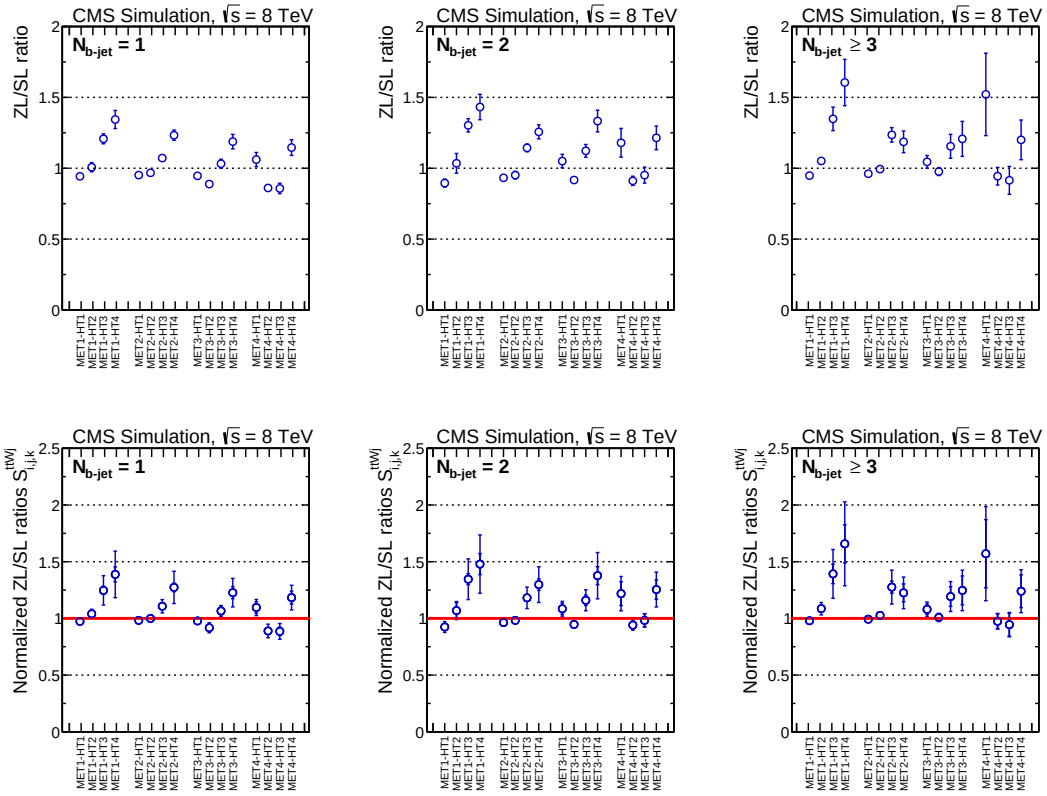


Figure 4.7: [top row] The ratio of ZL-to-SL $t\bar{t}$, W+jets, and single-top events in MC for each H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. The distributions are for $N_{b\mathrm{jet}} = 1$, 2, and ≥ 3 from left to right, respectively. The statistical uncertainties are included. [bottom row] The results from the ratio values plots in the top row divided by their corresponding average ratio values over all 48 analysis bins. The average is computed by integrating over all 3 dimensions of the ZL sample and dividing by the integral over all 3 dimensions of the SL sample. The distributions are for $N_{b\mathrm{jet}} = 1$, 2, and ≥ 3 from left to right, respectively. The inner (outer) error bars show the statistical (combined statistical and systematic) uncertainties. From [31].

average ratio values over all 48 analysis bins. The average is computed by integrating over all 3 dimensions of the ZL sample and dividing by the integral over all 3 dimensions of the SL sample. Deviations from one in the top row imply a shape discrepancy between the SL and ZL samples, and are accounted for in a bin-by-bin correction factor determined from the bottom row of Fig. 4.7.

Let indices i, j, k represent the three dimensions of $E_{\mathrm{T}}^{\mathrm{miss}}$, H_T, and N_{bjet}. The estimate of the $t\bar{t}$, W+jets, and single-top contribution to bin i, j, k of the ZL sample is given by

$$\mu_{ZL;\,i,j,k}^{ttWj} \;=\; S_{i,j,k}^{ttWj} \cdot R_{ZL/SL}^{ttWj} \cdot \mu_{SL;\,i,j,k}^{ttWj} \,, \tag{4.4}$$

where $\mu_{SL;\,i,j,k}^{ttWj}$ is the observed number of events (assuming 100% trigger efficiency) in the corresponding i-j-k bin of the SL sample, $R_{ZL/SL}^{ttWj}$ is the overall floating normalization parameter common to all bins, and the $S_{i,j,k}^{ttWj}$ are MC-derived bin-by-bin correction factors that account for the shape differences between the ZL and SL samples shown in the bottom row of Fig. 4.7.

The contribution of the T1bbbb signal is expected to be negligible in these bins, because the T1bbbb topology rarely contains high p_{T}-isolated leptons. Thus, the T1bbbb contribution, or "signal contamination," is not considered in the observable Poisson mean parameter for the $t\bar{t}$, W+jets, and single-top background. On the other hand, the T1tttt model contains four top quarks in the final state, where a high p_{T}-isolated lepton is typically found in the event. Thus, the T1tttt signal contamination is not negligible.

The likelihood model introduces the expected number of SUSY events in the SL region through the following relation:

$$\mu_{SL;\,i,j,k}^{SUSY} \;=\; I_{ZL}^{SUSY} \cdot \left(\mu_{SL;\,i,j,k}^{SUSY-MC} / I_{ZL}^{SUSY-MC}\right) \tag{4.5}$$

where $\mu^{SUSY}_{SL;i,j,k}$ is the expected number SUSY events in the i-j-k bin of the SL region, I^{SUSY}_{ZL} is the integrated number of SUSY events in the ZL sample and is allowed to float in the likelihood to determine the overall normalization of the SUSY signal, and $I^{SUSY-MC}_{ZL}$ is the integrated number of SUSY events in the ZL sample determined from MC.

The expected number of observed events, or the observable Poisson mean parameter, is given by

$$n_{SL;i,j,k} \;=\; \epsilon^{\rm trig}_{SL;i,j,k} \cdot (\mu^{ttwj}_{SL;i,j,k} + S^{\rm SUSY}_{{\rm SL};i,j,k} \cdot \mu^{\rm SUSY}_{{\rm SL};i,j,k}) \,, \tag{4.6}$$

where $\epsilon^{\rm trig}_{SL;i,j,k}$ is the bin-by-bin trigger efficiency correction measured in the SL sample described in Section 3.7.1, and $S^{\rm SUSY}_{{\rm SL};i,j,k}$ is a nuisance parameter. In the T1bbbb scenario, $\mu^{\rm SUSY}_{{\rm SL};i,j,k} = 0$.

4.2.2 QCD Background

In a technique similar to the one described above for the ttWj events, the QCD background is estimated by selecting a control sample, and then formulating an observable Poisson mean parameter for each of the analysis bins. By inverting the $\Delta\phi_N$ cut, the event sample is dominated by QCD events. This can be seen in Fig. 4.8, which shows the nominal selection except with $\Delta\phi_N < 4.0$ for distributions of $E^{\rm miss}_{\rm T}$, H_T, and $N_{b{\rm jet}}$. This region is referred to as the "low-delta phi" region, or LDP, and serves as a control sample for estimating the QCD background events. As mentioned in Section 3.5.1, the $\Delta\phi_N$ provides

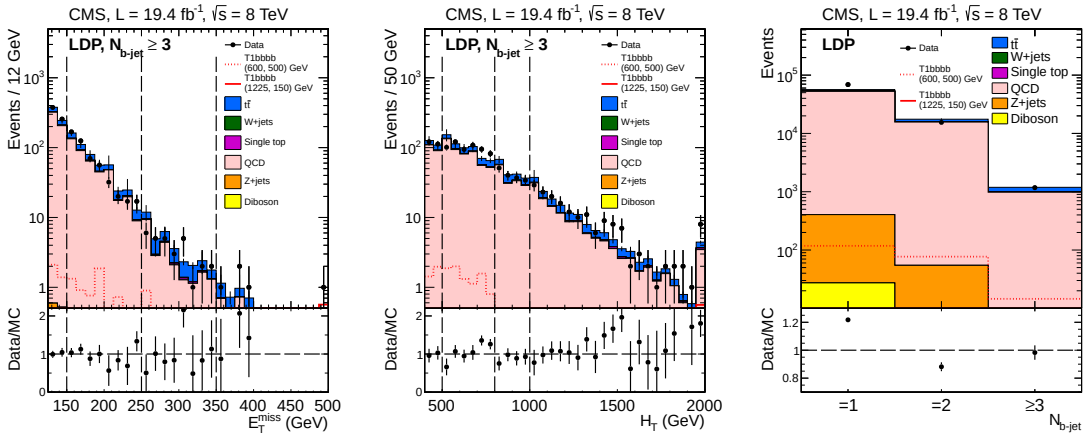


Figure 4.8: Data and MC distributions of $E_{\mathrm{T}}^{\mathrm{miss}}$ (left), H_T (middle), and $N_{b\mathrm{jet}}$ (right) for the LDP control sample. The accompanying ratio plots at the bottom show the ratio of data to simulated standard events. Dashed vertical lines indicate the $E_{\mathrm{T}}^{\mathrm{miss}}$ or H_T bin divisions. From [31].

a powerful discriminator for separating QCD and non-QCD events in a way that biases the $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in only a minimal way.

The estimation of QCD event yields in each 3D bin of the ZL sample is determined by applying a multiplicative scale factor to the corresponding bin of the LDP sample event yield. The scale factor is computed after subtracting contributions from other SM processes, such as ttWj and Z+jets events, in the LDP sample. The estimate of ttWj and Z+jets events are determined through yields from the likelihood fit for the corresponding ZL bin, and multiplied by the MC ratio of LDP to ZL events for that bin.

The top row of Fig. 4.9 shows the ratio of ZL-to-LDP events in MC for each H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. Each plot represents a different $N_{b\mathrm{jet}}$ bin, and the distributions are for $N_{b\mathrm{jet}} = 1$, 2, and ≥ 3 from left to right, respectively. Thus, there are 16 points shown for each plot, and there are three plots. The bottom row shows the ratios in the top plot divided by their

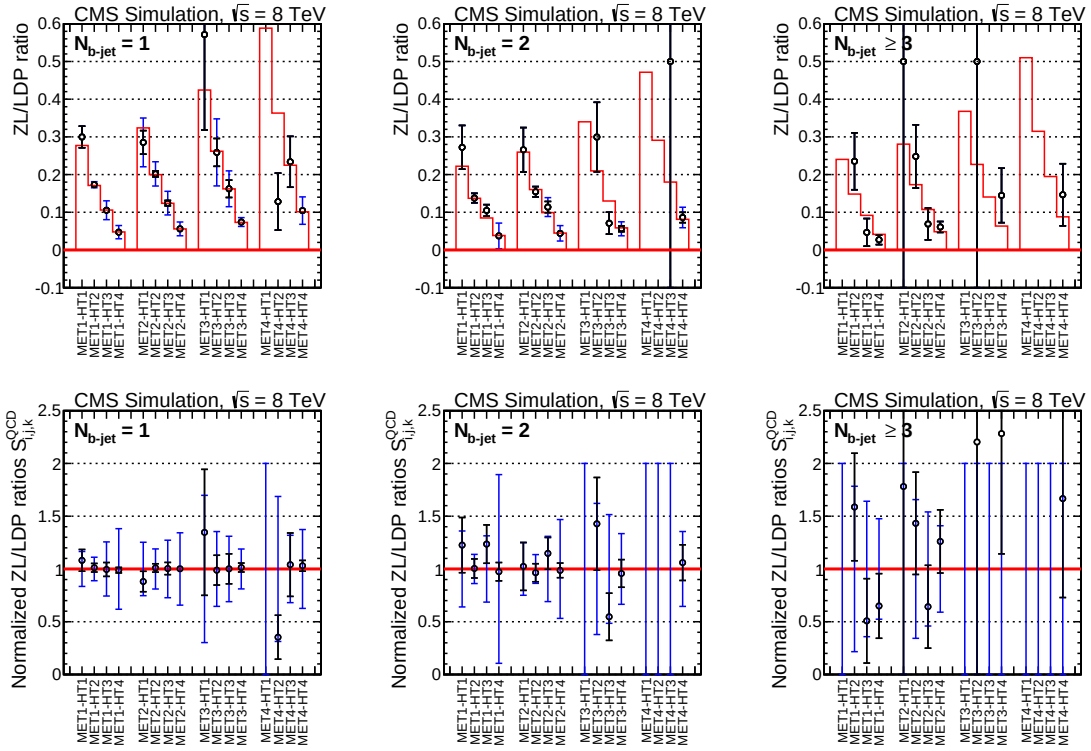


Figure 4.9: [top row] The ratio of ZL-to-LDP QCD events in MC for each H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. The distributions are for $N_{b\mathrm{jet}} = 1$, 2, and ≥ 3 from left to right, respectively. The inner error bars show the statistical uncertainties, while the outer error bars show the quadrature sum of statistical uncertainties and RMS values over several $\hat{p}_{\mathrm{T}}$ samples. The fitted parameterization results are shown in the histogram. [bottom row] The corresponding ratio divided by the parameterization from the top row. The inner (outer) error bars indicate the statistical (combined statistical and systematic) uncertainties. From Ref. [31].

average ratio value over all 48 analysis bins. The average is computed by integrating over all 3 dimensions of the ZL sample and dividing by the integral over all 3 dimensions of the SL sample. Deviations from one in the top row imply a shape discrepancy between the SL and LDP samples and are accounted for through bin-by-bin correction factors. Unlike the ttWj technique, the discrepancies are strongly correlated with H_T, and modestly so with $E_{\mathrm{T}}^{\mathrm{miss}}$ and $N_{b\mathrm{jet}}$. Thus, a correction factor exists for each dimension, giving three correction factors for each bin.

The estimate of the QCD contribution to bin i, j, k of the ZL sample is given by

$$\mu^{QCD}_{ZL;\,i,j,k} = S^{QCD}_{i,j,k} \cdot \left(K^{QCD}_{MET,i} \cdot K^{QCD}_{HT,j} \cdot K^{QCD}_{Nb,k} \right) \cdot \mu^{QCD}_{LDP;\,i,j,k} \,, \tag{4.7}$$

where $S^{QCD}_{i,j,k}$ are scale factors (see Ref. [31] for more details on how they are derived), the $K^{QCD}_{MET,i}$, $K^{QCD}_{HT,j}$, and $K^{QCD}_{Nb,k}$ terms are correction factors that account for the $E^{\rm miss}_{\rm T}$, H_T, and $N_{b\rm jet}$ dependence, respectively, and $\mu^{QCD}_{LDP;\,i,j,k}$ are the observed number events in the corresponding i-j-k bin of the LDP sample.

Since both T1bbbb and T1tttt signal contamination can be expected in the LDP region, the likelihood model introduces the expected number of SUSY events in the LDP region through the following relation:

$$\mu^{SUSY}_{LDP;\,i,j,k} \;=\; I^{SUSY}_{ZL} \cdot \left(\mu^{SUSY-MC}_{LDP;\,i,j,k} / I^{SUSY-MC}_{ZL} \right) \tag{4.8}$$

where $\mu^{SUSY}_{LDP;\,i,j,k}$ is the expected number SUSY events in the i-j-k bin of the LDP region, I^{SUSY}_{ZL} is the integrated number of SUSY events in the ZL sample and is allowed to float in the likelihood to determine the overall normalization of the SUSY signal, and $I^{SUSY-MC}_{ZL}$ is the integrated number of SUSY events in the ZL sample determined from MC.

Therefore, the expected number of observed events, or the observable Poisson mean parameter, in the QCD control sample is given by

$$n_{LDP;\,i,j,k} \;=\; \epsilon^{\text{trig}}_{ZL;i,j} \cdot \mu^{QCD}_{LDP;\,i,j,k} + \epsilon^{\text{trig}}_{SL;i,j} \cdot \left(\mu^{ttwj}_{LDP;\,i,j,k} + \mu^{Z\nu\nu}_{LDP;\,i,j,k} + S^{eff}_{LDP;\,i,j,k} \cdot \mu^{SUSY}_{LDP;\,i,j,k} \right) \tag{4.9}$$

where $\epsilon^{\text{trig}}_{ZL;i,j}$ and $\epsilon^{\text{trig}}_{SL;i,j}$ are the bin-by-bin trigger efficiency corrections for the ZL and SL samples, respectively, described in Section 3.7.1, $S^{eff}_{LDP;\,i,j,k}$ are bin-by-bin systematic uncertainty nuisance parameters constrained by lognormal PDFs in the likelihood, and $\mu^{ttwj}_{LDP;\,i,j,k}$ and $\mu^{Z\nu\nu}_{LDP;\,i,j,k}$ are the ttWj and Z $\rightarrow \nu\bar{\nu}$ contributions to the LDP sample.

4.3 Systematic Uncertainties

Systematic uncertainties arise from several sources and may associated with the:

- Jet energy scale (5-10%), where the size of the uncertainty depends on the jet p_{T} and η value
- Jet energy resolution (2%)
- Unclustered energy (1%)
- Pileup reweighting (3% - see Ref. [48])
- Anomalous $E_{\text{T}}^{\text{miss}}$ values due to event misreconstruction or beam noise (3%)
- Determination of the luminosity (4.4% - see Ref. [25])

- Trigger efficiency (below 2% - see Section 3.7.2)

- B-jet tagging efficiency (below 15%)

The number of events in the data control samples are typically small; thus, in this analysis, statistical uncertainties dominate the systematic uncertainties.

Chapter 5

Results

In this chapter, the global likelihood model and fit results are discussed. The pieces of the model, namely the formulation of the Poisson mean parameters for each of the background components, are described in Chapter 4. Because a significiant deviation of the number of data events is not observed from that predicted by the SM, the results of the fit are shown as 95% cross section upper limits on the T1bbbb and T1tttt models. A detailed description is given in Ref. [66], but the main points are summarized here.

A summary of all the Poisson mean parameters for the observerables in the likelihood are discussed in Chapter 4 and are included in Table 5.1.

Each of the observables in Table 5.1 represents the expected number of event counts in a given bin, and they are taken to be the means of the Poisson distributions. These distributions are the Poisson PDFs, and their products create the global likelihood function.

Signal search	$n_{ZL;\,i,j,k} = \epsilon^{\rm trig}_{ZL;i,j} \cdot \mu^{QCD}_{ZL;\,i,j,k}$
	$+ \epsilon^{\rm trig}_{SL;i,j} \cdot \left(\mu^{ttwj}_{ZL;\,i,j,k} + \mu^{Z\nu\nu}_{ZL;\,i,j,k} + S^{SUSY}_{ZL;\,i,j,k} \cdot \mu^{SUSY}_{ZL;\,i,j,k}\right)$
$t\bar{t}$ and W+jets control	$n_{SL;\,i,j,k} = \epsilon^{\rm trig}_{SL;i,j} \cdot (\mu^{ttwj}_{SL;\,i,j,k} + S^{\rm SUSY}_{\rm SL;\,i,j,k} \cdot \mu^{\rm SUSY}_{\rm SL;\,i,j,k})$
QCD control	$n_{LDP;\,i,j,k} = \epsilon^{\rm trig}_{ZL;i,j} \cdot \mu^{QCD}_{LDP;\,i,j,k}$
	$+ \epsilon^{\rm trig}_{SL;i,j} \cdot \left(\mu^{ttwj}_{LDP;\,i,j,k} + \mu^{Z\nu\nu}_{LDP;\,i,j,k} + S^{eff}_{LDP;\,i,j,k} \cdot \mu^{SUSY}_{LDP;\,i,j,k}\right)$
Z $\to \ell^+\ell^-$ control	$n_{zll;\,i,j,k} = \mu^{zll}_{zll;\,i,j,k}/P_{ll}$

Table 5.1: Poisson mean parameters for the observables in the likelihood. From Ref. [66].

Thus the numbers of signal and background events are determined using a global likelihood fit, which incorporates all appropriate statistical and systematic uncertainties.

The parameters in Table 5.1 give the means of the Poisson distributions. In other words, the Poisson PDF gives the probability to observe N events given a mean n. The likelihood is evaluated at the observed Ns. The HT1-MET4 bin of Fig. 4.1 is kinematically unlikely, because events with large $E_{\rm T}^{\rm miss}$ tend to have large H_T, so few events populate it. This bin is excluded from the analysis, and the number of effective observables drops to 165.

Since the n parameters incorporate contribution from SUSY, a search for SUSY can be incorporated by determining the agreement between n and N as the SUSY contribution is increased. The measure of how much SUSY is being observed is accounted for in the I^{SUSY}_{ZL} parameter. If this parameter is significantly larger than zero, the data is consistent with a signal.

I^{SUSY}_{ZL} was not observed to be significantly larger than zero in either the T1bbbb or T1tttt scenario, so upper limits on the cross sections are measured for these models using

the asymptotic CL_s [49, 59] method. The efficiencies of the ZL signal selection are shown in Fig. 5.1. In the plots, the T1bbbb scenario is is referred to as the $\tilde{g}\tilde{g} \to 2 \times b\bar{b}\tilde{\chi}_1^0$ process, and the T1tttt scenario is referred to as the $\tilde{g}\tilde{g} \to 2 \times t\bar{t}\tilde{\chi}_1^0$ process. The upper limits on the T1bbbb cross sections computed at 95% confidence level (CL) are shown in Fig. 5.2. The solid black lines show the exclusion using a reference NLO + next-to-leading logarithm (NLL) cross section [9, 53, 52, 8, 7] for gluino pair production, and the dashed black lines show ± 1 standard deviation theoretical uncertainty [51] on this limit. The dashed red lines show the expected limit with their corresponding ± 1 standard deviation experimental uncertainties.

As seen in Fig. 5.1, the T1bbbb (T1tttt) efficiency is roughly 60% (25%) over most of the m_{gluino}-m_{LSP} phase space, but drops to 40% to 20% (5% to 15%) close to the diagonal intersecting the $m_{\text{LSP}} = 0$ axis at around $m_{\text{gluino}} = 400\,\text{GeV}$ (550 GeV) or for gluino masses below about 550 GeV (680 GeV). The rapid drop in efficiency is due to the Q value (kinetic energy that is released from the decay of a particle at rest) becoming smaller as the m_{LSP} value gets closer to the m_{gluino}. The smaller Q value results in a slower momentum for the LSP particle, which means that the $E_{\text{T}}^{\text{miss}}$ in the signal event is lower. From Fig. 5.2, it is seen that for small values of m_{LSP}, m_{gluino} is excluded up to around 1170 GeV (1020 GeV). Mass ranges up to 650 GeV (350 GeV) are excluded for m_{LSP}. The results at the time of this writing are among the most stringent bounds obtained.

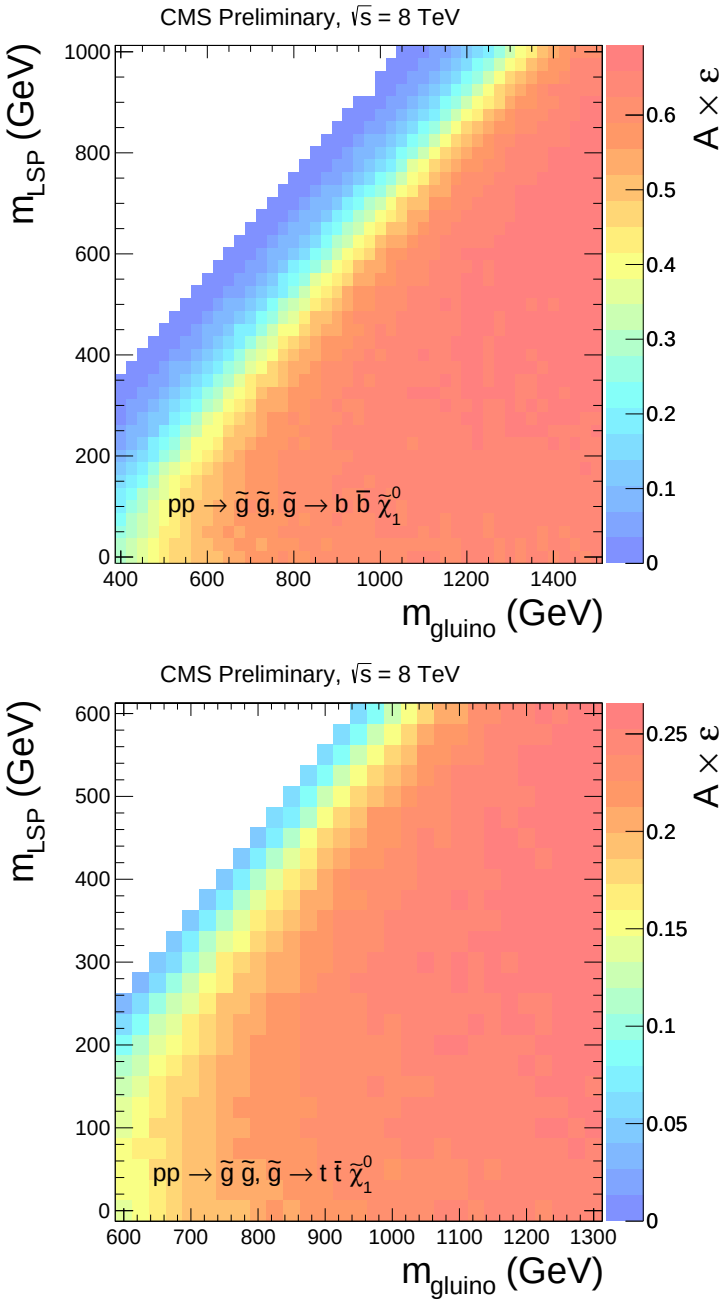


Figure 5.1: Signal (ZL) selection efficiencies for the T1bbbb (top) and T1tttt (bottom) scenarios. From [31].

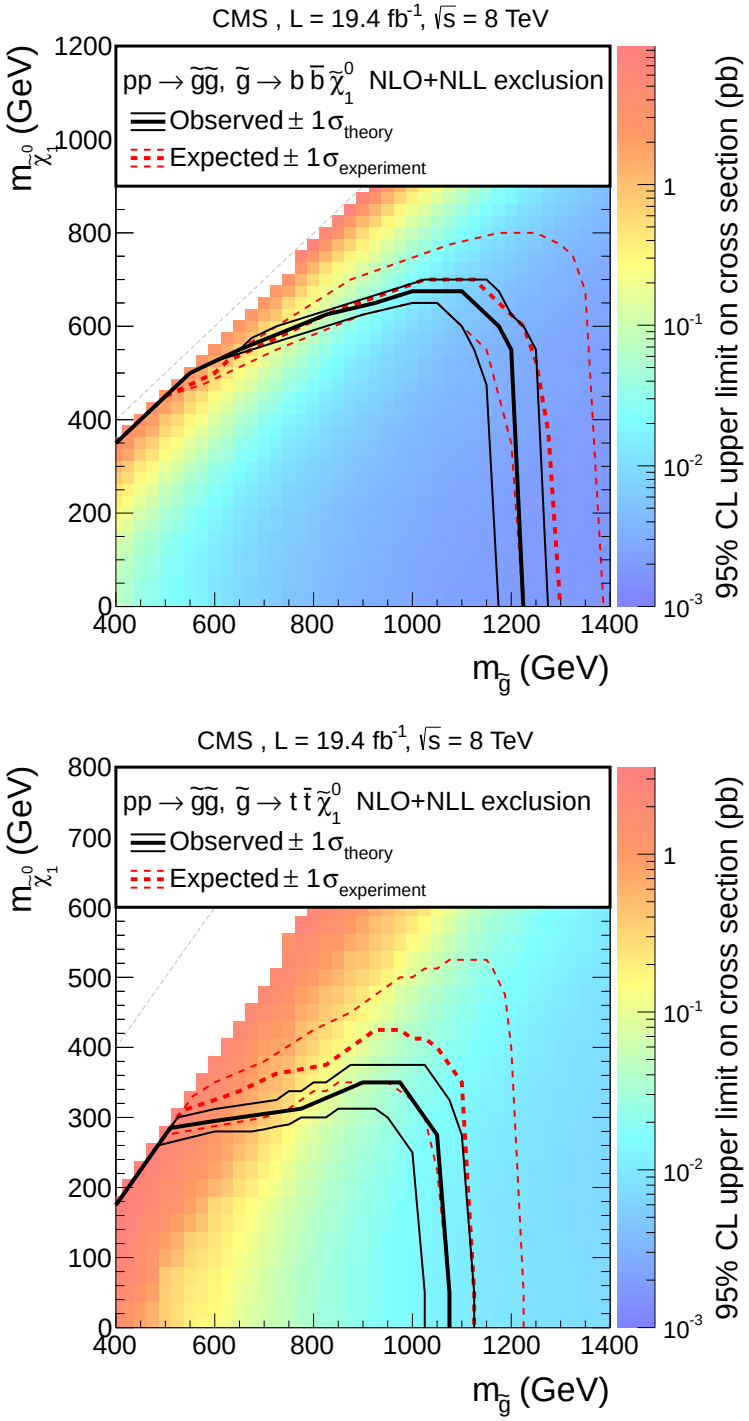


Figure 5.2: The upper limits on the T1bbbb cross sections computed at 95% confidence level (CL) for the T1bbbb (top) and T1tttt scenariors (bottom). The solid black lines show the exclusion using a reference NLO+NLL cross section for gluino pair production, and the dashed black lines show ± 1 standard deviation theoritical uncertainty on this limit. The dashed red lines show the expected limit with their corresponding ± 1 standard deviation experimental uncertainties. From [31].

Chapter 6

Summary

A search is presented for physics beyond the standard model based on events with large missing transverse energy, at least three jets, and at least one identified b jet. This thesis discussed a brief overview of the analysis procedures performed in Ref. [31], with particular emphasis on the trigger efficiency measurements and photon control sample selection. The study is based on a sample of $19\,\mathrm{fb}^{-1}$ collected at 8 TeV with the CMS detector at the Large Hadron Collider in 2012. The main standard model backgrounds come from $t\bar{t}$, $W+\mathrm{jets}$, $\mathrm{Z} \to \nu\bar{\nu}$, and single-top processes. Backgrounds are estimated using data control samples, and 165 mutually exclusive observables in the dimension of H_T, $E_{\mathrm{T}}^{\mathrm{miss}}$, and b-jet multiplicity are simultaneously analyzed in a global likelihood fit.

The data are found to be consistent with standard model processes, and the results are interpreted in the context of simplified models of supersymmetry. Upper limits on the production cross sections of the T1bbbb and T1tttt simplified models are presented, which

are supersymmetry scenarios of gluino pair production where each gluino decays into a bottom quark-antiquark pair or a top quark-antiquark pair, respectively, plus an undetected particle. Gluino masses up to 1170 GeV are excluded for the T1bbbb scenario and up to 1020 GeV for the T1tttt scenario, at 95% confidence level.

Bibliography

[1] Amnon Harel. Forward jet quality criteria. *CMS Analysis Note*, AN-12-067, 2010.

[2] E. Aprile and S. Profumo. EDITORIAL: Focus on Dark Matter and Particle Physics. *New Journal of Physics*, 11(10):105002, October 2009.

[3] S. Ask, M.A. Parker, T. Sandoval, M.E. Shea, and W.J. Stirling. Using gamma+jets Production to Calibrate the Standard Model Z(nunu)+jets Background to New Physics Processes at the LHC. *J. High Energy Phys.*, 1110:058, 2011.

[4] Riccardo Barbieri and Duccio Pappadopulo. S-particles at their naturalness limits. *J. High Energy Phys.*, 10:061, 2009.

[5] D Barney. A pedagogical introduction to the cms electromagnetic calorimeter. Technical Report CMS-CR-1998-004, CERN, Geneva, 1998.

[6] CMS Collabration Bayatian et al. *CMS Physics Technical Design Report Volume I: Detector Performance and Software*. Technical Design Report CMS. CERN, Geneva, 2006.

[7] Wim Beenakker, Silja Brensing, Michael Krämer, Anna Kulesza, Eric Laenen, Leszek Motyka, and Irene Niessen. Squark and gluino hadroproduction. *Int. J. Mod. Phys. A*, 26:2637, 2011.

[8] Beenakker, Wim and Brensing, Silja and Krämer, Michael and Kulesza, Anna and Laenen, Eric and Niessen, Irene. Soft-gluon resummation for squark and gluino hadroproduction. *J. High Energy Phys.*, 12:041, 2009.

[9] Beenakker, Wim and Höpker, R. and Spira, M. and Zerwas, P. M. Squark and gluino production at hadron colliders. *Nucl. Phys. B*, 492:51, 1997.

[10] J Beringer et al. Review of particle physics. *Phys. Rev. D.*, 86:010001, 2012.

[11] J Beringer et al. Review of Particle Physics, 2012-2013. Review of Particle Properties. *Phys. Rev. D*, 86(1):010001, 2012.

[12] L. Borrello et al. Sensor design for the cms silicon strip tracker. Technical Report CMS-NOTE-2003-020, CERN, Geneva, Aug 2003.

[13] Buras, A.J., Ellis, J., Gaillard, M.K., Nanopoulous, D.V. Aspects of the Grand Unification of Strong, Weak and Electromagnetic Interactions. *Nucl. Phys. B*, 135:66–92, 1978.

[14] M. Cacciari, G. P. Salam, and G. Soyez. The anti-k_t jet clustering algorithm. *J. High Energy Phys.*, 04:063, 2008.

[15] M. S. Carena and H. E. Haber. Higgs Boson Theory and Phenomenology. *Prog.Part.Nucl.Phys*, 50:63–152, 2003.

[16] cfA team. "cfA PF2PAT recipe". `http://cmssw.cvs.cern.ch/cgi-bin/cmssw.cgi/UserCode/pbgeff/cmssw/Workspace/ConfigurableAnalysis/python/runningPatOnFly_cfg.py?revision=1.41`, 2012.

[17] Serguei Chatrchyan et al. Missing transverse energy performance of the CMS detector. *JINST*, 6:P09001, 2011.

[18] Serguei Chatrchyan et al. Search for New Physics with Jets and Missing Transverse Momentum in pp collisions at $\sqrt{s} = 7$ TeV. *J. High Energy Phys.*, 1108:155, 2011.

[19] CMS Collaboration. Algorithms for b jet identification in cms. CMS Physics Analysis Summary CMS-PAS-BTV-09-001, CERN, Geneva, 2009.

[20] CMS Collaboration. "HCAL PFG Task List (May 2009 - October 2009)". `https://twiki.cern.ch/twiki/bin/viewauth/CMS/HcalPromTaskList`, 2009.

[21] CMS Collaboration. Particle-flow event reconstruction in cms and performance for jets, taus and met. *CMS PAS PFT-09-001*, 2009.

[22] CMS Collaboration. Comparison of the fast simulation of CMS with the first LHC data. CMS Detector Performance Summary CMS-DP-2010-039, CERN, Geneva, 2010.

[23] CMS Collaboration. "Photon reconstruction and identification". CMS Physics Analysis Summary CMS-PAS-EGM-10-005, CERN, 2010.

[24] CMS Collaboration. "CMS Luminosity - Public Results". `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`, 2012.

[25] CMS Collaboration. Cms luminosity based on pixel cluster counting – summer 2012 update. CMS Physics Analysis Summary CMS-PAS-LUM-12-001, CERN, Geneva, 2012.

[26] CMS Collaboration. "HLT config browser". `http://j2eeps.cern.ch/cms-project-confdb-hltdev/browser/`, 2012.

[27] CMS Collaboration. Search for supersymmetry in events with b-quark jets and missing transverse energy in pp Collisions at 7 TeV. *CMS PAS SUS-12-003*, 2012.

[28] CMS Collaboration. "SUSY RA2b - b jets + jets + MET: 2011 Analysis". https://twiki.cern.ch/twiki/bin/view/CMS/SusyRA2BJets2011, 2012.

[29] CMS Collaboration. "SusyPhotonID". https://twiki.cern.ch/twiki/bin/view/CMS/SusyPhotonID, 2012.

[30] CMS Collaboration. "Tools for conversion rejection (electron ID) and electron vetoing (photon ID)". https://twiki.cern.ch/twiki/bin/view/CMS/ConversionTools, 2012.

[31] CMS Collaboration. Search for gluino-mediated bottom- and top-squark production in pp collisions at 8 tev. *CMS PAS SUS-12-012*, 2013.

[32] CMS Collaboration. "Standard Model Cross Sections for CMS at 8 TeV". https://twiki.cern.ch/twiki/bin/view/CMS/StandardModelCrossSectionsat8TeV, 2013.

[33] CMS E/gamma POG. E/gamma 2012 analysis supporting material–electrons. https://twiki.cern.ch/twiki/bin/viewauth/CMS/Eg2012AnalysesSupportingMaterial#Electrons, 2012.

[34] CMS Muon POG. Baseline muon selections for 2012 data (cmssw 52x and above). https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId#The2012Data, 2012.

[35] ATLAS Collaboration. Observation of an excess of events in the search for the standard model higgs boson with the atlas detector at the lhc. Technical Report ATLAS-CONF-2012-093, CERN, Geneva, Jul 2012.

[36] CMS Collaboration. Data-driven estimation of the invisible z background to the susy met plus jets search. Technical Report CMS-PAS-SUS-08-002, CERN, 2009. Geneva, Jan 2009.

[37] CMS Collaboration. Determination of jet energy calibration and transverse momentum resolution in cms. *JINST*, 6:P11002, 2011.

[38] CMS Collaboration. Observation of a new boson with a mass near 125 GeV. Technical Report CMS-PAS-HIG-12-020, CERN, Geneva, 2012.

[39] S. Dimopoulos and G.F. Giudice. Naturalness constraints in supersymmetric theories with nonuniversal soft terms. *Phys. Lett. B*, 357:573, 1995.

[40] G. R. Farrar and P. Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Phys. Lett. B*, 76:575, 1978.

[41] Pierre Fayet. Supergauge invariant extension of the Higgs mechanism and a model for the electron and its neutrino. *Nucl. Phys. B*, 90:104, 1975.

[42] G. Abbiendi, et al., The LEP Working Group for Higgs Boson Searches. Search for the Standard Model Higgs Boson at LEP. *Phys.Lett.*, B565:61–75, 2003.

[43] Georgi, H., Glashow, S.L. Unity of All Elementary Particle Forces. *Phys. Rev. Lett.*, 32:438–441, 1974.

[44] Yu.A. Golfand and E.P. Likhtman. Extension of the algebra of Poincaré group generators and violation of p invariance. *JETP Lett.*, 13:323, 1971.

[45] J.D. Griffiths. *Introduction to Elementary Particle Physics*. Wiley VCH, 08 2008.

[46] Harel, A. and Schieferdecker, P. Jet identification. `https://twiki.cern.ch/twiki/bin/view/CMS/JetID`, 2010.

[47] Hatakeyama, K. and Schoefbeck, R. Met optional filters. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETOptionalFilters`, 2012.

[48] Hildreth, M. Estimating systematic errors due to pileup modeling. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupSystematicErrors`, 2011.

[49] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instr. and Meth. A*, 434:435, 1999.

[50] Vardan Khachatryan et al. Search for Supersymmetry in pp Collisions at 7 TeV in Events with Jets and Missing Transverse Energy. *Phys.Lett.*, B698:196–218, 2011.

[51] Michael Krämer, Anna Kulesza, Robin van der Leeuw, Michelangelo Mangano, Sanjay Padhi, et al. Supersymmetry production cross sections in pp collisions at $\sqrt{s} = 7$ TeV. 2012.

[52] Kulesza, Anna and Motyka, L. Soft gluon resummation for the production of gluino-gluino and squark-antisquark pairs at the LHC. *Phys. Rev. D*, 80:095004, 2009.

[53] Kulesza, Anna and Motyka, L. Threshold resummation for squark-antisquark and gluino-pair production at the LHC. *Phys. Rev. Lett.*, 102:111802, 2009.

[54] A. Neveu and J.H. Schwarz. Factorizable dual model of pions. *Nucl. Phys. B*, 31:86, 1971.

[55] Hans Peter Nilles. Supersymmetry, supergravity and particle physics. *Phys. Rep.*, 110:1, 1984.

[56] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, et al. "Implications of CTEQ global analysis for collider observables". *Phys. Rev.*, D78:013004, 2008.

[57] Michele Papucci, Joshua T. Ruderman, and Andreas Weiler. Natural SUSY endures. *J. High Energy Phys.*, 09:035, 2012.

[58] Pierre Ramond. Dual theory for free fermions. *Phys. Rev. D*, 3:2415, 1971.

[59] A. L. Read. Presentation of search results: the CL_s technique. *J. Phys. G*, 28:2693, 2002.

[60] ROOT Development Team. "ROOT — A Data Analysis Framework". `http://root.cern.ch/drupal/`.

[61] S. Agostinelli et al. "GEANT4: A Simulation toolkit". *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[62] S. Eidelamn et al. "Passage of Particles Through Matter". *Physics Letters*, B592:1, 2004.

[63] T. Sjöstrand, S. Mrenna, and P. Skands. PYTHIA 6.4 physics and manual. *J. High Energy Phys.*, 05:026, 2006.

[64] Sprenger, D. and others. "Tags for PAT Layer 1 production V12 (52X/53X) [r18]". `https://twiki.cern.ch/twiki/bin/view/CMS/SusyPatLayer1DefV12?rev=18`, 2012.

[65] The RA2b Group. Search for new physics in events with b-jets and missing transverse energy with the full 2011 data sample. *CMS Analysis Note*, AN-11-409, 2011.

[66] The RA2b Group. Search for gluino-mediated bottom- and top-squark production in pp collisions at 8 TeV. *CMS Analysis Note*, AN-12-081, 2012.

[67] D.V. Volkov and V.P. Akulov. Possible universal neutrino interaction. *JETP Lett.*, 16:438, 1972.

[68] J. Wess and B. Zumino. A Lagrangian model invariant under supergauge transformations. *Phys. Lett. B*, 49:52, 1974.

[69] J. Wess and B. Zumino. Supergauge transformations in four dimensions. *Nucl. Phys. B*, 70:39, 1974.