University of California
Santa Barbara

A Search for Supersymmetry in Events with Jets and Missing Momentum Using Data Collected by the CMS Detector at 13 TeV

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

Doctor of Philosophy
in
Physics

by

Sam Daniel Mullin

Committee in charge:

Professor Joseph Incandela, Chair
Professor David Berenstein
Professor Benjamin Monreal

June 2016

The Dissertation of Sam Daniel Mullin is approved.

Professor David Berenstein

Professor Benjamin Monreal

Professor Joseph Incandela, Committee Chair

May 2016

A Search for Supersymmetry in Events with Jets and Missing Momentum Using Data Collected by the CMS Detector at 13 TeV

by

Sam Daniel Mullin

Curriculum Vitæ

Sam Daniel Mullin

Education

2016	Ph.D. in Physics (Expected), University of California, Santa Barbara.
2010	B.S. in Physics and B.S. in Applied Mathematics, University of California, Riverside

Publications

CMS Collaboration, "A Search for Scalar Top Quark Production and Decay to All Hadronic Final States in pp Collisions at $\sqrt{s} = 8$ TeV", CMS-PAS-SUS-13-023, 2015, `http://cds.cern.ch/record/2044441`.

CMS Collaboration, "Search for direct production of top squark pairs decaying to all-hadronic final states in pp collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-SUS-16-007, 2016, `http://cds.cern.ch/record/2141543`.

Abstract

A Search for Supersymmetry in Events with Jets and Missing Momentum Using Data Collected by the CMS Detector at 13 TeV

by

Sam Daniel Mullin

As successful as the Standard Model of particle physics has been it still has several major shortcomings which range from unanswered theoretical questions to a lack of any explanation for observed phenomena such as dark matter. One proposed theory for physics beyond the Standard Model which provides solutions for some of these issues is supersymmetry. This dissertation presents a search for supersymmetry using 2.3 fb^{-1} of proton-proton collision data. This data was collected at a center-of-mass energy of 13 TeV by the CMS detector at the LHC during 2015. This search focuses on top squark pair production where the produced stops both decay to an all hadronic final state. These decays are characterized by multiple jets and missing transverse momentum. A baseline search region is defined to be sensitive to signal processes which occur at rates many orders of magnitude lower than Standard Model processes. The sensitivity to various signal models is improved by dividing this baseline region into distinct categories. Events with an unreconstructed lepton from leptonic W boson decays constitute the primary background. There are also significant contributions from events where a Z boson decays invisibly to neutrinos especially in bins with higher missing transverse momentum. Events with multijet production where one jet has been severely mis-measured as well as those with a pair of top quarks and an invisibly decaying Z boson also have a small presence in the search region.

The contributions from these processes to the search region is estimated using data control regions. No statistically significant deviations from the predicted background yields are observed. The results are interpreted in terms of exclusion limits using the Simplified Model Spectrum framework. Stop and neutralino masses are probed up to 780 GeV and 260 GeV respectively.

Contents

Chapter 1

Introduction

The drive to explain the world can be seen in ancient mythologies where gods were created as the personifications of natural phenomena. Over time these attempts have evolved to incorporate empirical evidence from observations and experiments. This allows various claims to be tested and kept only when they are in agreement with the evidence. Modern technology has reached a point where the observation of progressively more fundamental pieces of nature has become a reality. During the second half of the 20th century, significant strides were made on this front both experimentally and theoretically. The the Standard Model of particle physics is the result of this progress. This framework successfully explained every experimental observation at the time as well as predicting several additional massive particles. Although they were not predicted by the Standard Model, observations of neutrino mixing were easily incorporated into it. With the announcement of the discovery of the Higgs boson in July 2012, the final predicted piece had been found.

Recent astronomical observations, however, have shown that this is not the end of the story. In fact, it is only a small fraction of the mass of the universe.

Numerous experimental and theoretical endeavors are underway to discover what else could be out there. This document presents one such search for evidence of physical processes that are beyond the Standard Model. Data collected by the Compact Muon Solenoid (CMS) detector during the 2015 data taking period is used to look for signs of new heavy particles [1] which are predicted by many of the theoretical extensions of the Standard Model.

Chapter 1 gives an overview of the theoretical background and experimental setup, beginning with a discussion of the current state of high energy physics in section 1.1. Specifically, section 1.1.1 gives a brief description of the Standard Model and its fundamental particles. Section 1.1.2 then highlights several of shortcomings of the Standard Model and one of the proposed extensions is described in section 1.1.3. The experimental setup and reconstruction of collected data are summarized in sections 1.2 and 1.3 respectively. The search strategy is described in chapter 2. This includes the datasets and objects that are used as well as the event selection requirements which are necessary to distinguish events containing new physics from ones with known processes. A detailed description of the methods used to estimate the number of remaining background events is in chapter 3. Finally, chapter 4 presents the results of this search including interpreting them in the context of new physics models.

1.1 Theoretical Motivation

1.1.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics [2–4] is a framework describing the fundamental particles and forces which are the basic building blocks of nature.

force	mediator	mass
strong	gluon (g)	0
electromagnetic	photon (γ)	0
weak	$W^{\pm}$, Z	80.4, 91.2 GeV

Table 1.1: The fundamental forces in the SM and their associated gauge bosons [5].

The particles can be divided into two primary groups: bosons with integer spin and fermions with half-integer spin. The SM successfully incorporates three of the four fundamental forces: the strong, electromagnetic, and weak forces. The remaining force, gravity, is entirely negligible at the energy scales currently accessed by collider experiments. Therefore, its absence is not of immediate concern and it will be ignored for the majority of this discussion. Each force is mediated by one or more gauge bosons as shown in table 1.1. Particles with color and electric charge interact through the strong and electromagnetic forces respectively. All particles interact through the weak force.

The fermions can be divided into two distinct groups: quarks and leptons, as shown in tables 1.2 and 1.3 respectively, both of which come in three "generations". Each generation of quarks or leptons is identical except that the masses of each subsequent one increase considerably. Each fermion also has an anti-particle, denoted as $\bar{f}$, which is oppositely signed but otherwise identical to it. Each quark comes in three colors and thus interacts through the strong force while all leptons are color-neutral. Single quarks are never observed, rather, they are always bound as baryons (qqq or $\bar{q}\bar{q}\bar{q}$) or mesons ($q\bar{q}$).

There is one final boson that has recently been discovered, the Higgs, with a mass of 125.7 GeV [6,7]. Its role in the SM will be discussed in more detail at the end of this section.

particle	symbol	mass
up	u	2.3 MeV
down	d	4.8 MeV
charm	c	1.3 GeV
strange	s	95 MeV
top	t	173 GeV
bottom	b	4.2 GeV

Table 1.2: Quarks grouped by generation [5]. The upper quark in each box has an electric charge of $+\frac{2}{3}$ while the lower one has an electric charge of $-\frac{1}{3}$.

Quantum Field Theory

So far this discussion has only been a very brief account of what has been observed in nature. A more formal treatment of particles and their observed interactions is formulated in terms of a relativistic quantum field theory. Within this formalism, every particle is an excitation of some underlying field and interactions are described by a Lagrangian. These underlying fields cannot be measured, therefore changing a field's phase should have no observable impact on the Lagrangian. This is known as gauge invariance. While it is fairly simple to write down a Lagrangian that is invariant under a global phase transformation, it seems odd that an identical transformation must be applied to all of space-time simultaneously, including parts which are not in communication with each other, in order to keep the Lagrangian invariant. What happens if invariance under a local phase transformation, i.e. $\psi \rightarrow e^{i\alpha(x)}\psi$ where $\alpha(x)$ is some function of space-time, is required? It turns out that requiring the Lagrangian to be invariant under a local gauge transformation results in the gauge bosons and their couplings to particles naturally emerging from the theory.

particle	symbol	mass
electron	e	0.51 MeV
electron neutrino	ν_e	≈ 0
muon	μ	106 MeV
muon neutrino	ν_μ	≈ 0
tau	τ	1.8 GeV
tau neutrino	ν_τ	≈ 0

Table 1.3: Leptons grouped by generation [5]. The neutrino masses are listed as ≈ 0 since, although neutrino mixing requires them to have mass, as yet only differences between and an upper limit on their masses have been measured. Electrons, muons, and taus all have an electric charge of +1 while all the neutrinos are neutral.

To illustrate this, consider a simple Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi.$$

Although it is clearly invariant under a global phase transition, under a local one $\mathcal{L}$ picks up an extra term from the derivative of α: $i(\partial_\mu\alpha(x))e^{i\alpha(x)}\psi$, thus breaking its invariance. To enforce local gauge invariance, a covariant derivative ($D_\mu = \partial_\mu - ieA_\mu$) must be introduced as a replacement for every ∂_μ. In the covariant derivative, A_μ is a gauge field that transforms as $A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\alpha$ under a local gauge transformation. An additional term ($-\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2A_\mu A^\mu$ where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$) is also introduced in order to treat A_μ as a gauge particle. However, the mass term is not invariant under a local gauge transformation therefore this new field must be massless. Thus a locally gauge invariant version of the above Lagrangian is:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi + e\bar{\psi}\gamma^\mu A_\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}.$$

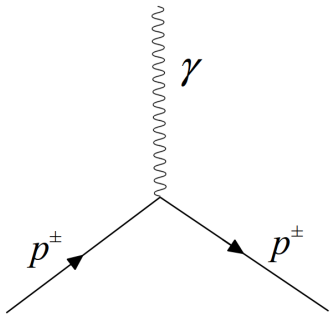

Figure 1.1: The QED vertex for the interaction of a photon with a charged particle.

This Lagrangian accurately describes quantum electrodynamics (QED) with the massless field A_μ identified as the photon. The vertex for the interaction of a photon with a charged particle is shown in figure 1.1. Since $e^{i\alpha}$ is a 1×1 unitary matrix, this symmetry is called $U(1)$ gauge invariance.

For the strong interactions of quantum chromodynamics (QCD), ψ from the above example must be replaced with a triplet for each quark flavor, $\bar{q} = (\bar{q}_r, \bar{q}_b, \bar{q}_g)$, to account for the three color charges. A local gauge transformation can then be written as $q \rightarrow e^{i\alpha_a(x)T_a}q$ where the T_a are the eight generators of $SU(3)$. This results in the Lagrangian:

$$\mathcal{L}_{\text{QCD}} = i\bar{q}\gamma^\mu\partial_\mu q - g(\bar{q}\gamma^\mu T_a q)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a,$$

with $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g f_{abc} G^b_\mu G^c_\nu$. This has the expected terms for free quarks and gluons as well as quark-gluon interactions. However, unlike the QED Lagrangian above, it also contains terms for self-interactions between gluons. This reflects the gluons' own color charge which is the source of quark confinement where quarks are confined to color-neutral combinations as well as the hadronization of any quarks that become separated. The interaction vertices for quarks and gluons are shown in figure 1.2.

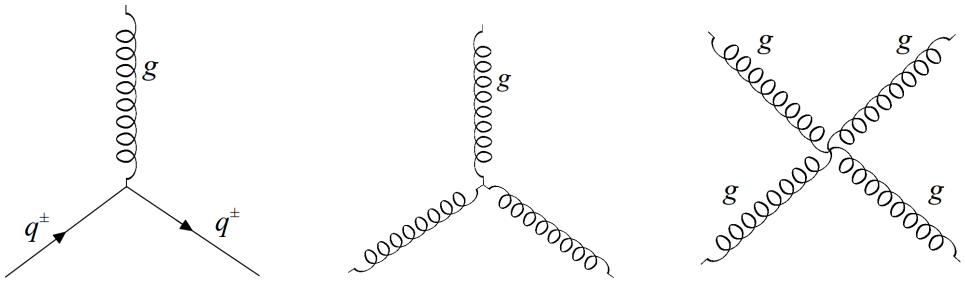

Figure 1.2: The QCD vertex for the interaction of a quark and a gluon (left) and the three (middle) and four (right) gluon vertices.

Spontaneous Symmetry Breaking

Although requiring the Lagrangian to remain invariant under a local gauge transformation works for introducing massless photons and gluons, the observed non-zero masses of the $W^{\pm}$ and Z bosons present a problem. Clearly the same approach will not be sufficient to establish a Lagrangian for electroweak interactions. Luckily this problem can be resolved with a process called spontaneous symmetry breaking where the initial symmetry of a Lagrangian is broken by the transformation to a ground state.

To illustrate spontaneous symmetry breaking, consider a locally invariant Lagrangian with a complex scalar field $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$ of the form:

$$\mathcal{L}_h = (D_\mu \phi)^* (D^\mu \phi) + \mu^2 \phi^* \phi - \lambda (\phi^* \phi)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}.$$

Here μ and λ are both real constants and the potential component of $\mathcal{L}_h$ has a circle of minima at $\phi_1^2 + \phi_2^2 = \mu^2/\lambda = v^2$. In order to use perturbative calculations, $\mathcal{L}_h$ must be rewritten in terms of new fields with a minimum at zero: $\eta = \phi_1 - v$ and $\xi = \phi_2$, where the minimum $\phi_1 = v$, $\phi_2 = 0$ is used. $\mathcal{L}_h$ can then be rewritten

as:

$$\mathcal{L}_h = \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2\right] + \left[\frac{1}{2}(\partial_\mu \xi)(\partial^\mu \xi)\right] + \left[\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu\right] + \left[evA_\mu \partial^\mu \xi + \text{other interaction terms}\right]$$

where η is now a massive particle, ξ is a massless Goldstone boson, and A_μ is the massive gauge field needed for an electroweak boson. To eliminate the Goldstone boson and the problematic $A_\mu \partial^\mu \xi$ term, the local gauge transformation

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) = \frac{1}{\sqrt{2}}(v + \eta + i\xi) \rightarrow \phi' = \frac{1}{\sqrt{2}}(v + \eta + i\xi)e^{i\alpha(x)}$$

can be carefully chosen so that ϕ' is real. The above $\mathcal{L}_h$ then reduces to:

$$\mathcal{L}_h = \left[\frac{1}{2}(\partial_\mu \eta)(\partial^\mu \eta) - \mu^2 \eta^2\right] + \left[\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2 v^2 A_\mu A^\mu\right] + \left[\text{interaction terms}\right]$$

where all that remains is a massive particle which can be identified as the Higgs and a massive gauge field.

Extending this "Higgs mechanism" to the $SU(2) \times U(1)$ electroweak gauge group results in three massive bosons ($W^\pm$, Z) and one massless boson (γ) as well as a Higgs field (h). A few example electroweak and Higgs vertices are in figure 1.3. This same Higgs field also couples to all of the quarks and leptons, giving them mass in a locally gauge invariant way. Thus the entire observed Standard Model can be described by the $SU(3) \times SU(2) \times U(1)$ gauge group.

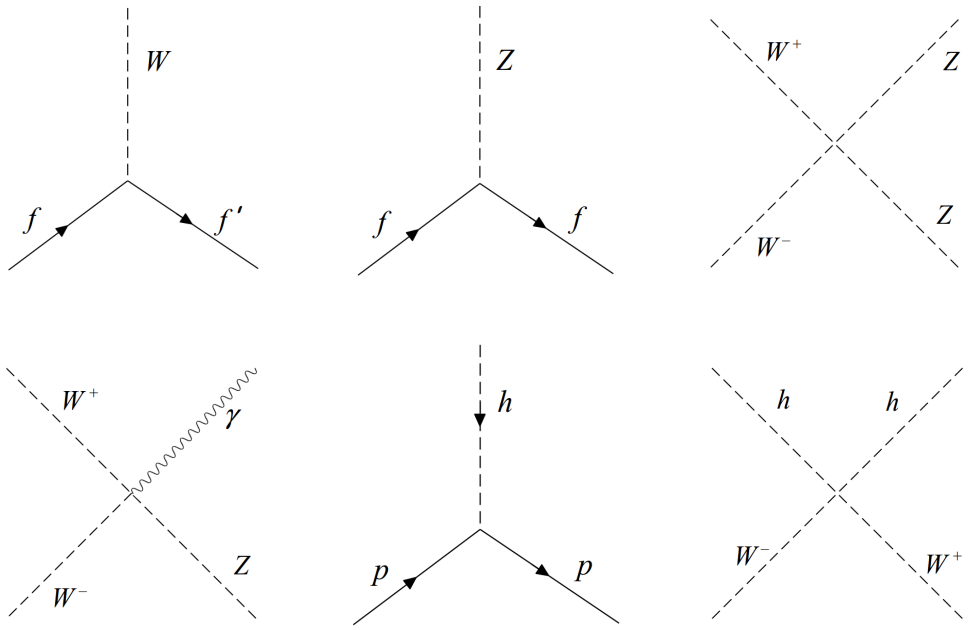

Figure 1.3: A few example electroweak and Higgs interaction vertices.

1.1.2 Shortcomings of the Standard Model

As well-tested as the Standard Model has been over the past few decades, there are still several major shortcomings of the theory [8]. These range from unanswered theoretical questions to convenient cancellations to a complete lack of any explanation for some observed phenomena. Many theories have been proposed for physics beyond the standard model (BSM). While these theories vary considerably in their underlying physics, they must effectively reduce to the current standard model at the lower energy scales which have already been well tested as well as providing solutions to at least some of the outstanding issues.

One of the more glaring deficiencies of the standard model is the lack of any feasible candidate for dark matter or dark energy, which are estimated to make up about 25% and 70% of the energy of the universe respectively [5]. For example, experimental evidence from the observed rotation curves of galaxies and clusters

of galaxies suggests that a "dark halo" of non-luminous, weakly interacting matter must exist. Although some of these observations can be resolved by modifying the theory of gravity at the large distances involved, so far any attempts to make these theories relativistic have been unsuccessful. Dark matter candidates need to be stable, at least on a cosmological timescale, and interact with electromagnetic radiation very weakly if at all. They also need to be sufficiently heavy such that they were not relativistic when galaxies began to form. One potential candidate is weakly interacting massive particles (WIMPs), which are expected to have masses on the order of 10 GeV to a few TeV and roughly weak strength cross sections.

With the recent discovery of the Higgs boson at a mass of 125.7 GeV the so-called "hierarchy problem" has ceased to be hypothetical. This problem is caused by higher order corrections to the Higgs mass from loop diagrams that are quadratically divergent. The Higgs mass is calculated as:

$$M_H^2 = (M_H^2)_{bare} + \mathcal{O}(\lambda, g^2, h^2)\Lambda^2$$

where λ, g^2, and h^2 are the couplings between the Higgs and other particles. Λ is the cutoff scale in the theory and presumably around the Plank scale ($\Lambda \sim 10^{19}$ GeV). Thus the expected scale for the Higgs mass is on the order of Λ. Given the observation of a fairly low mass Higgs, the bare value squared must cancel with these corrections to more than 30 orders of magnitude.

There are also several shortcomings of a more theoretical nature. These include the significant number of arbitrary parameters which can only be determined experimentally and the inability to integrate gravity into the SM. It is also desirable to unite the strong and electroweak forces into a single force at some higher energy. These considerations, along with patterns in the observed "fundamental"

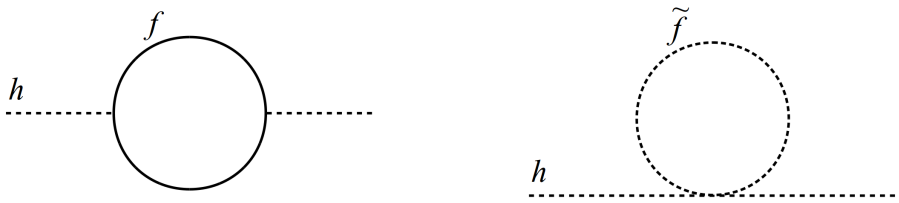

Figure 1.4: One-loop corrections to M_H^2 from a fermion (left) and its superpartner (right).

particle spectrum (e.g. three generations of fermions), suggest that there is some underlying structure.

One particularly popular BSM theory is supersymmetry, which is the topic of the next section. The supersymmetric particles provide a natural way of canceling the divergent corrections to the Higgs mass. Moreover, the lightest of these particles may be stable and thus a potential candidate for dark matter.

1.1.3 Supersymmetry

Supersymmetry (SUSY) [9] is a symmetry which relates bosons and fermions such that every particle is associated to another particle whose spin differs by a half-integer. These related pairs of particles are referred to as superpartners. All properties other than spin are identical between each pair of superpartners. The new SUSY particles, collectively called "sparticles" and denoted with a tilde ($\widetilde{p}$), have the same mass and couplings as their SM counterparts. Therefore the magnitude of their contributions to the mass of the Higgs boson are identical and have opposite signs. This results in all of the corrections that depend on Λ canceling perfectly and thus the hierarchy problem is solved. An example pair of one-loop diagrams that contribute to the Higgs mass are shown in figure 1.4.

As appealing as SUSY is, there is a glaring problem with its existence in

nature: if sparticles exist with the same mass as their SM counterparts they would have already been detected. Since this is clearly not the case, SUSY must be a spontaneously broken symmetry. However, in order to retain SUSY as a viable solution to the hierarchy problem, it cannot be broken too badly. The most alarming corrections to M_H^2 have the form:

$$\Delta M_H^2 \propto (\lambda_S - |\lambda_f|^2)\Lambda^2 + ...$$

In order to maintain the cancellations of these quadratically diverging corrections after SUSY breaking the equivalence between the couplings of a particle and its associated superpartner must be maintained. This is referred to as "soft" SUSY breaking.

There are still other corrections to M_H^2 of the form:

$$\Delta M_H^2 \propto m_{\text{soft}}^2 \left[\lambda \ln\left(\frac{\Lambda}{m_{\text{soft}}}\right) + ...\right]$$

where m_{soft} is the mass scale associated with soft SUSY breaking. These terms only scale with $\ln(\Lambda)$ so their contributions are manageable as long as m_{soft} is not too large. This requires the sparticles to not be too massive since the mass splitting between a particle and its superpartner is determined by m_{soft}. The masses of at least the lightest sparticles should not be heavier than the TeV scale in order to avoid overly suspicious cancellations.

So far there is nothing to restrict SUSY interactions that violate baryon (B) and lepton (L) number conservation. This is problematic since experimental evidence such as the lack of any sign of proton decays has restricted any non-conservation to very low levels. One solution is to simply refuse to include any

such terms in the SUSY Lagrangian. However, since B and L conservation are already required by the SM a supposedly more fundamental theory that does not also enforce this is not ideal. A more appealing option is to add a new symmetry called "R-parity" or "matter parity" which prohibits the inclusion of B and L violating terms.

R-parity is a discrete symmetry in which all SM particles have even R-parity (+1) and all SUSY particles have odd R-parity (-1). Every interaction is required to conserve the product of every particles' R-parity. Equivalently every allowed vertex must have an even number of sparticles. Therefore sparticles must be produced in pairs. This also means that the lightest supersymmetric particle (LSP) must be completely stable since there will by definition exist no lighter sparticle into which it can decay. If the LSP is also electrically neutral and thus only interacts weakly with ordinary matter it is a promising candidate for dark matter.

SUSY Particle Spectrum

Although there is a wide variety of possible SUSY models, it is useful to initially restrict the discussion to the minimum requirements for a consistent theory. This is called the Minimal Supersymmetric Standard Model (MSSM). Since none of the currently observed SM particles can be superpartners of each other the MSSM must at least double the current particle list. The superpartners of the SM fermions are named by adding an "s" to the beginning of the SM name while the superpartners of the SM bosons are named by modifying the SM name so that it ends with "-ino". This is demonstrated in tables 1.4 and 1.5 respectively.

A single superpartner might be expected for the SM Higgs boson, however, a single Higgs cannot couple to both up-type and down-type quarks in a supersym-

fermion		sfermion	
quark	q	squark	$\widetilde{q}$
top	t	stop	$\widetilde{t}$
bottom	b	sbottom	$\widetilde{b}$
lepton	ℓ	slepton	$\widetilde{\ell}$
electron	e	selectron	$\widetilde{e}$
neutrino	ν	sneutrino	$\widetilde{\nu}$

Table 1.4: Example supersymmetric partners of the SM fermions called sfermions.

metric theory. A minimum of five Higgs mass eigenstates are actually required: three neutral ones (h^0, H^0, A^0) and two charged ones ($H^{\pm}$). The lightest neutral Higgs (h^0) corresponds to the SM Higgs.

None of the sparticles are required to also be mass eigenstates other than the gluino. Due to its color charge the gluino cannot mix with any other sparticles. Electroweak symmetry breaking in particular results in the higgsinos and electroweak gauginos mixing considerably. The neutral (charged) higgsinos combine with the neutral (charged) gauginos to form four mass eigenstates called neutralinos: $\widetilde{\chi}_1^0$, $\widetilde{\chi}_2^0$, $\widetilde{\chi}_3^0$, and $\widetilde{\chi}_4^0$ (charginos: $\widetilde{\chi}_1^{\pm}$ and $\widetilde{\chi}_2^{\pm}$), which are labeled in order of

boson		bosino	
gluon	g	gluino	$\widetilde{g}$
W boson	W^0,$W^{\pm}$	wino	$\widetilde{W}^0$, $\widetilde{W}^{\pm}$
B boson	B^0	bino	$\widetilde{B}^0$
Z boson	Z^0	zino	$\widetilde{Z}^0$
photon	γ	photino	$\widetilde{\gamma}$
Higgs	h	higgsino	$\widetilde{h}$

Table 1.5: Supersymmetric partners of the SM bosons called bosinos.

sparticle	example decays				
$\widetilde{q}$	$q\widetilde{g}$	$q\widetilde{\chi}_i^0$	$q'\widetilde{\chi}_i^\pm$		
$\widetilde{g}$	$q\widetilde{q}$				
$\widetilde{\ell}^\pm$	$\ell^\pm\widetilde{\chi}_i^0$	$\nu\widetilde{\chi}_i^\pm$			
$\widetilde{\nu}$	$\nu\widetilde{\chi}_i^\pm$	$\ell^\pm\widetilde{\chi}_i^0$			
$\widetilde{\chi}_i^0$	$Z\widetilde{\chi}_j^0$	$W^\pm\widetilde{\chi}_j^\mp$	$h^0\widetilde{\chi}_j^0$	$\ell^\pm\widetilde{\ell}^\mp$	$\nu\widetilde{\nu}$
$\widetilde{\chi}_i^\pm$	$W^\pm\widetilde{\chi}_j^0$	$Z\widetilde{\chi}_1^\pm$	$h^0\widetilde{\chi}_1^\pm$	$\ell^\pm\widetilde{\nu}$	$\nu\widetilde{\ell}^\pm$

Table 1.6: Examples of some of the more likely sparticle decays.

increasing mass. Any mixing of sfermions is usually assumed to be minimal.

Squarks will usually decay to a quark and a gluino if the gluino is lighter than a given squark. They can also decay to a quark plus either a neutralino or chargino. Gluinos can only decay to a (quark,squark) pair. Sleptons can decay to a chargino/neutralino plus a lepton. Since the charginos and neutralinos all contain mixtures of the higgsinos and electroweak gauginos they can decay to both (boson,bosino) and (fermion,sfermion) pairs. Examples of allowed decay modes are given in table 1.6. Three or four body decays can also occur although at much lower rates.

Although different SUSY models with various parameters result in very diverse predictions for the sparticle mass spectrum, there are a few common reasonable assumptions relevant for this analysis which are worth highlighting. The lightest neutralino is often assumed to be the LSP. This makes it is a good dark matter candidate since it is only weakly interacting and can have a fairly large mass. The stop and sbottom are expected to be the lightest of the squarks. This a good sign for SUSY's ability to resolve the hierarchy problem since the top quark contributes the largest corrections to the Higgs mass.

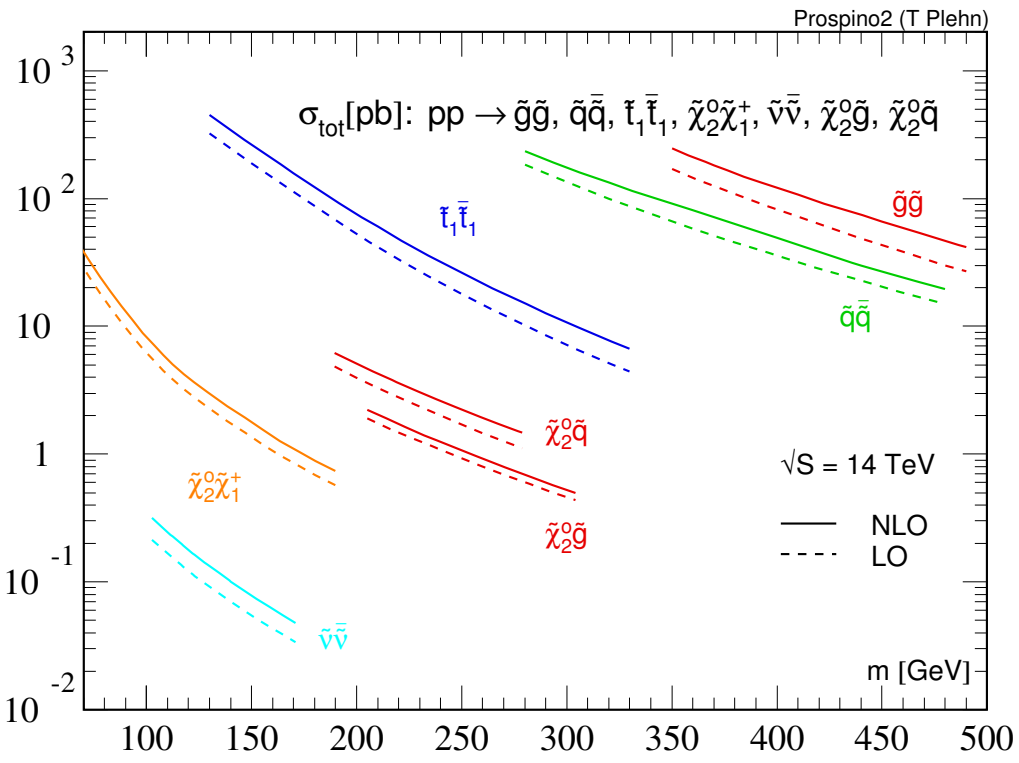

Figure 1.5: Pair production cross sections for sparticles at the LHC at the design energy of 14 TeV [10].

Searching for Direct Stop Production

Under the assumption of R-parity conservation all sparticles must be produced in pairs. At the LHC, squark and gluino pair production is expected to dominate as can be seen in figure 1.5. Since the analysis that will be described in the following chapters is a search for stop pair production the remainder of this discussion will focus on the decays of stop/anti-stop pairs.

Although a stop can decay in a variety of ways depending on the specifics of the sparticle mass spectrum, this can be reduced to just a few observable outcomes with an additional assumption: the stop is assumed to be the lightest of the squarks as well as lighter than the gluino. Combing this with the assumption

that the lightest neutralino is the LSP, the relevant decay modes are $\widetilde{t} \to t\widetilde{\chi}_1^0$ and $\widetilde{t} \to b\widetilde{\chi}^{\pm}$. The top in the first case decays as $t \to bW^{\pm}$ and the chargino in the second case will most likely decay as $\widetilde{\chi}^{\pm} \to W^{\pm}\widetilde{\chi}_1^0$. Both of these decay chains result in a penultimate state of $bW^{\pm}\widetilde{\chi}_1^0$ so a single search can be designed with both final states in mind.

The Simplified Model Spectra (SMS) framework [11–13] is used to design a search for these decay modes. This framework is a way of reducing the number of new particles and parameters in a way that still reproduces kinematic observables. Interpreting search results in the context of simplified models allows them to be extended to other more complicated models that have a similar final state. The simplified models relevant for this search are shown in figure 1.6. The first, labeled T2tt, assumes that both stops decay directly to a top and LSP ($\widetilde{t} \to t\widetilde{\chi}_1^0$). Second, T2bW assumes that both stops decay to a chargino which then decays to a W boson and LSP ($\widetilde{t} \to b\widetilde{\chi}^{\pm} \to bW^{\pm}\widetilde{\chi}_1^0$). A third model, T2tb, covers the mixed case where one stop decays directly to a top and LSP and the other decays through an intermediate chargino. Due to technical difficulties with the production of the T2bW samples it is not currently considered in the interpretation of this analysis.

1.2 The CMS Experiment at CERN

1.2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [14], which is run by the European Organization for Nuclear Research (CERN), is currently the largest particle collider in the world. Located on the Swiss-French boarder outside Geneva Switzerland, it is housed in a 27 km circular tunnel which is about 100 m underground. Al-

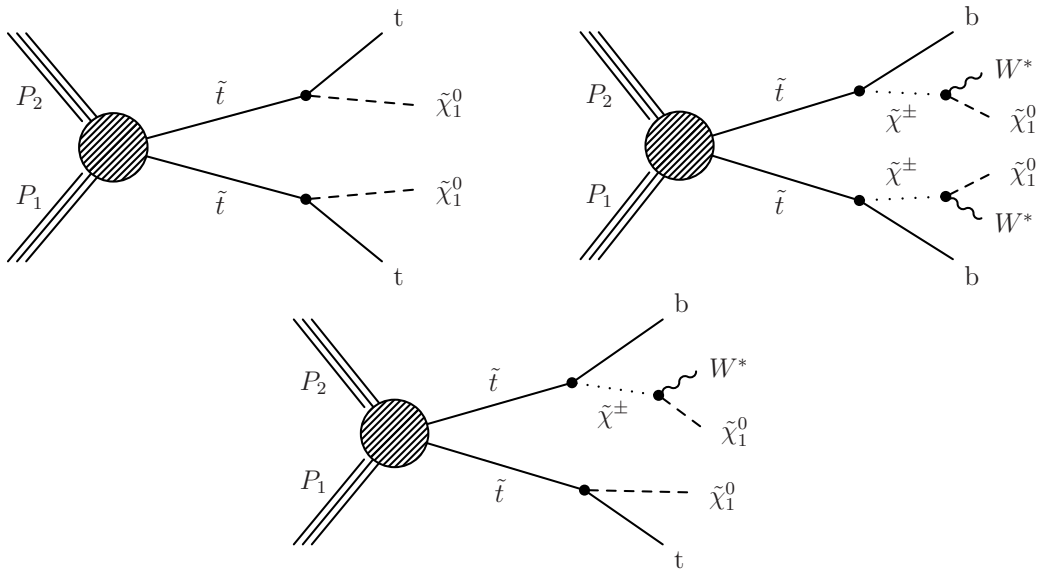

Figure 1.6: Diagrams of the simplified models for direct stop pair production that are targeted by this search. They are T2tt (top left) where each stop decays to a top and LSP, T2bW (top right) where both stops decay through an intermediate chargino, and T2tb (bottom) where one stop decays to a top and LSP and the other decays through an intermediate chargino.

though it has been designed to reach a center-of-mass energy of $\sqrt{s} = 14$ TeV for proton collisions and 1.15 PeV for heavy ion collisions, it is currently running at $\sqrt{s} = 13$ TeV.

To reach this energy, protons from ionized hydrogen gas are accelerated with the following series of accelerators (also shown in figure 1.7): the Linac2 (50 MeV), the Proton Synchrotron Booster (1.4 GeV), the Proton Synchrotron (25 GeV), and the Super Proton Synchrotron (450 GeV). Finally, two beams are injected into the LHC to circle in opposite directions. Here they are brought to their final energies of 6.5 TeV each for a total center of mass energy of 13 TeV. Over 1500 superconducting magnets, which are cooled below 2 K and have magnetic fields over 8 T, are used to control the beams and bring them together at each of the

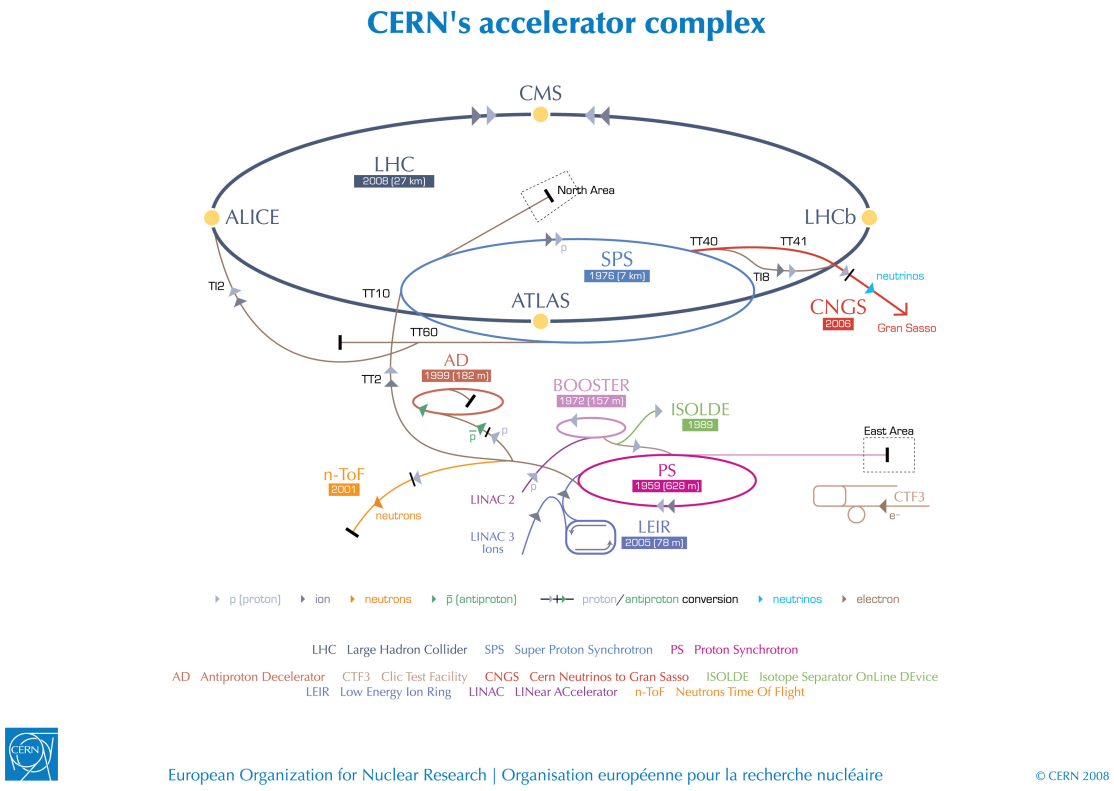

Figure 1.7: Depiction of the accelerator complex at CERN [15].

four detectors: ALICE, ATLAS, CMS, and LHCb.

The LHC is designed to be able to reach a maximum instantaneous luminosity (L) of 10^{34} cm^{-2}s^{-1}. Anti-particle beams, which have been used for particle/anti-particle colliders in the past, cannot reach the high beam intensity required for this luminosity. Therefore the LHC uses particles for both beams. The particle beams are arranged into bunches so that collisions only take place in discrete time intervals. Given this scheme, there are three adjustable parameters that affect the instantaneous luminosity: the number of particles in each bunch (N_b), the total number of bunches in the beam (n_b), and the amplitude of the transverse motion of a particle within the beam at the interaction point (which is controlled by the function β^*). These parameters are related as $L \propto N_b^2 n_b / \beta^*$. There are physical

limitations on their optimal achievable values as follows.

The particle density per bunch is limited by the beam-beam interactions of the particles when they collide. Thus N_b is limited by the maximum tolerable linear shift from these interactions. The minimum value of β^* is limited by the size of the apertures of the magnets. The primary limiting factors for n_b are the minimum bunch spacing of 25 ns and the need for the occasional longer gaps between the bunches that are necessary to reset electronics and dump the beam.

In addition to the above physical limitations on N_b, n_b, and β^*, optimizing their values for high luminosity also has consequences when reconstructing the events for analysis. Increasing N_b and decreasing β^* result in more collisions per bunch crossing. This makes reconstructing the primary (i.e. highest energy) interaction more difficult. Increasing n_b makes differentiating between subsequent crossings more difficult since more particles from the previous crossing are still in the detector.

Particle interactions cause the instantaneous luminosity to decay over time. In order to maximize the total integrated luminosity over the full data-taking period the length of time a given pair of beams is kept in the accelerator must be optimized against the expected time required to refill the LHC and ramp the beams to their final energy. During the proton-proton collisions in 2015, the total integrated luminosity recorded by CMS is 3.8 fb^{-1}, and shown in Figure 1.8.

1.2.2 The Compact Muon Solenoid Detector

The Compact Muon Solenoid (CMS) detector [17] is located 100 m underground just outside of Cessy, France. One of the two general purpose detectors on the LHC, it is 21.6 m long, 14.6 m in diameter, and weighs 14,000 tons.

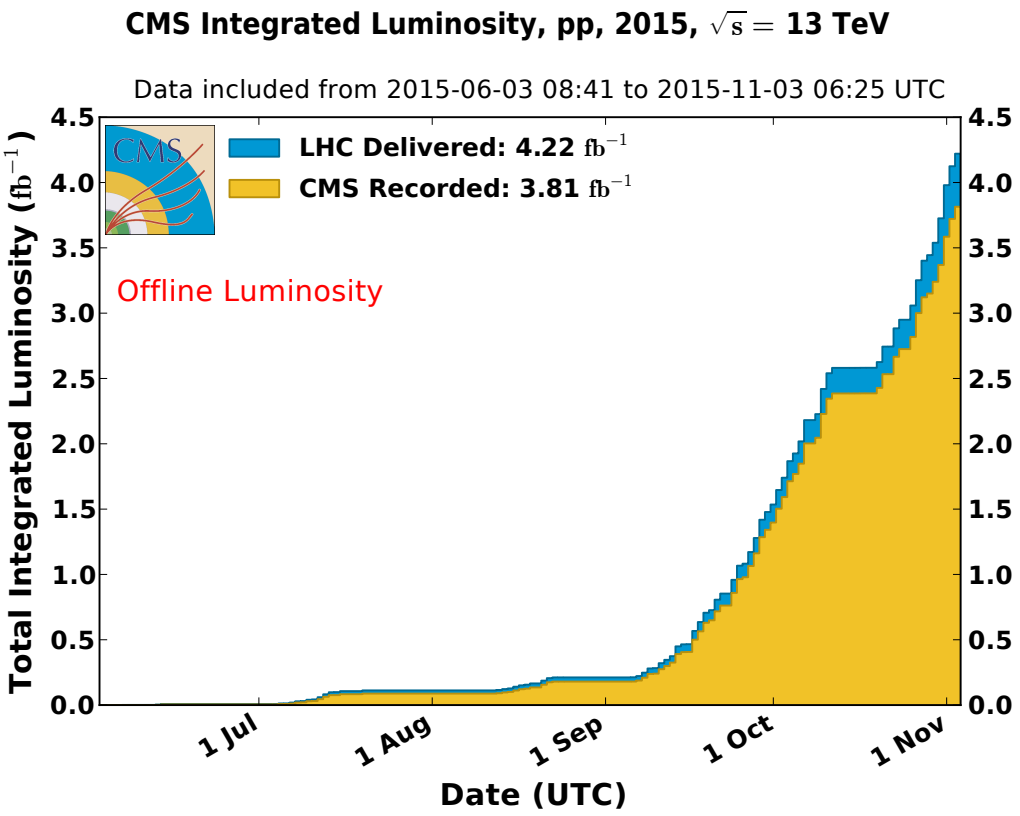

Figure 1.8: Total cumulative integrated luminosity from proton-proton collisions delivered by the LHC (blue) and collected by CMS (yellow) in 2015 [16].

Figure 1.9 has a photo of the detector with the endcaps and barrel sections separated. The most notable feature of CMS is the 3.8 T superconducting solenoid that is 12.5 m long, 220 tons, and capable of storing up to 2.6 GJ. The magnet's 6 m inner diameter has space for the tracker, electromagnetic calorimeter, and hadronic calorimeter. The magnet's iron return yolk houses the muon systems. The large bending field of the magnet allows the momentum of charged particles to be precisely measured by the tracker. Housing the electromagnetic and hadronic calorimeters inside the magnet increases the precision of energy measurements by reducing energy losses before particles reach the calorimeters. A slice of the

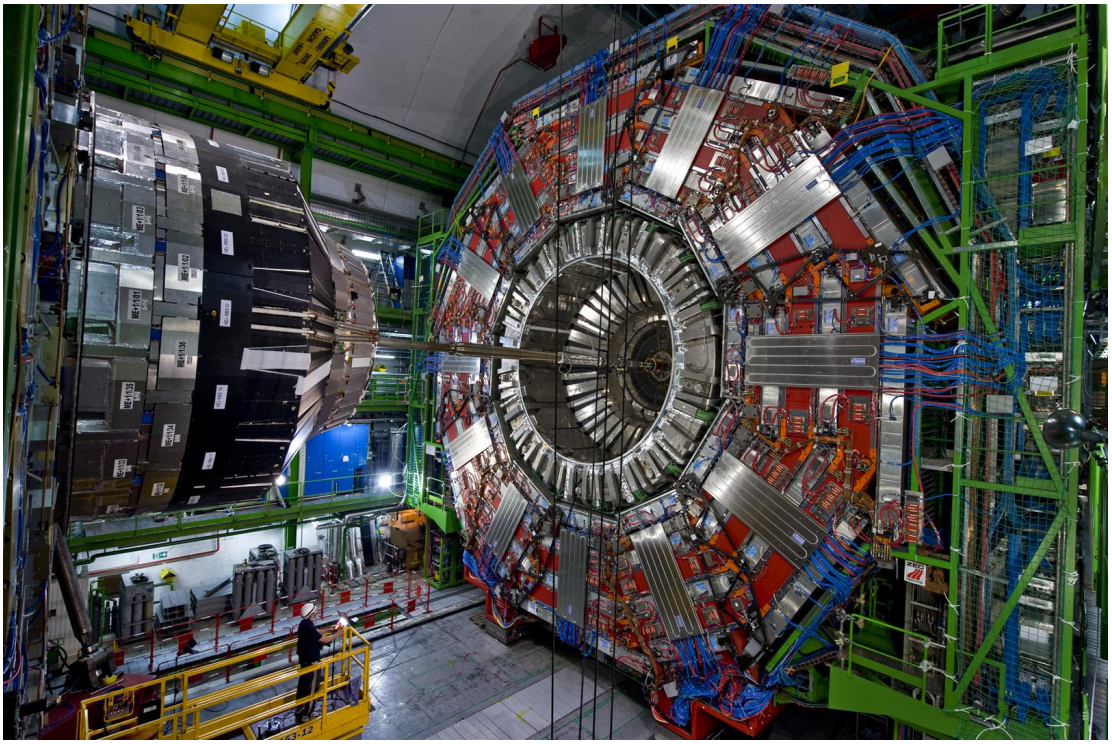

Figure 1.9: The CMS detector before the endcaps were closed in 2008. The endcap is on the left, the barrel is on the right, and the beam pipe has been installed [18].

detector with example particles is shown in figure 1.10.

CMS uses a right handed coordinate system with the nominal collision point at the origin to define positions within the detector. The x, y, and z-axes point toward the center of the LHC ring, vertically upward, and counter-clockwise along the beam respectively. In the xy-plane the azimuthal angle (ϕ) is measured from the x-axis and the radius (r) is the distance from the origin. The polar angle (θ) is measured from the z-axis. Pseudorapidity, which is defined as $\eta = \ln[\tan(\theta/2)]$, is used instead of θ since it is roughly invariant for Lorentz boosts along the z-axis. Thus positions in the detector can be given with (r, η, ϕ). For convenience, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is often used to denote the angular distance between objects in (ϕ, η). All sub-detectors have a cylindrical barrel component that is centered

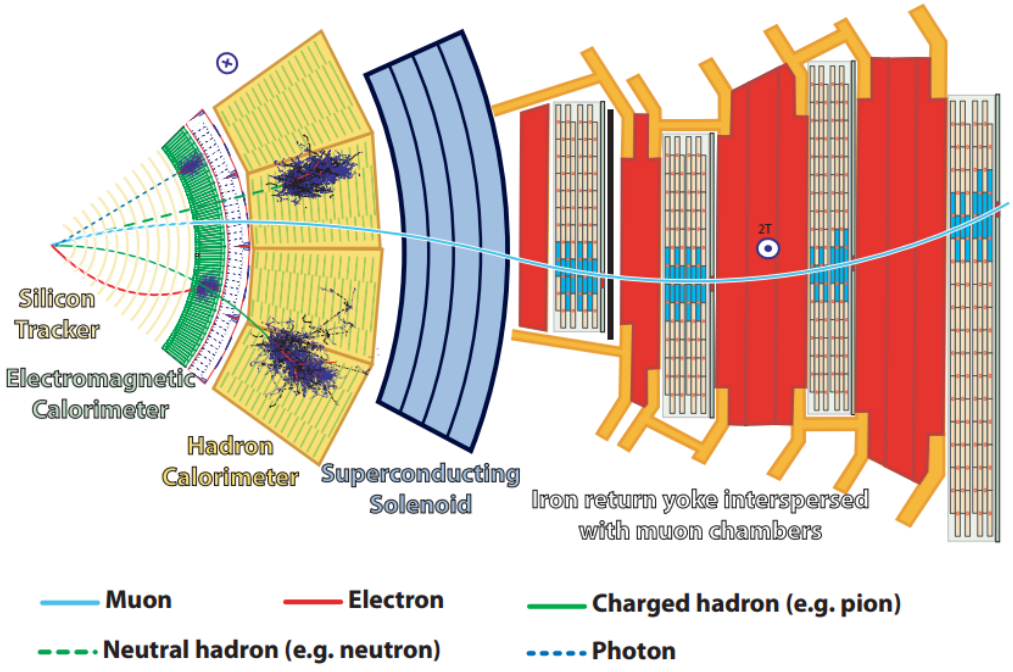

Figure 1.10: Slice of the CMS detector that demonstrates how several different types of particles interact [19].

along the beam for lower $|\eta|$ and two endcaps in the x-y plane on either side for higher $|\eta|$. Many quantities only use a quantity's component in the transverse plane since the beamline makes complete coverage in η impossible and the actual interaction energy of hadronic collisions is not knowable. The most widely used example of this is the transverse momentum of a particle, denoted as p_{T}.

The high luminosity delivered by the LHC means that there will be on the order of one thousand charged particles in every event. Most of these are from interactions other than the one of interest. These additional particles are referred to as pile-up (PU). Particles from collisions in a given bunch crossing that are not from the primary interaction are in-time PU. Particles remaining in the detector from previous bunch crossings are out-of-time PU. While all sub-detectors are sus-

ceptible to in-time PU to various degrees, elements of the detector with response times greater than the bunch spacing of 25 ns are also affected by out-of-time pileup. The impact of PU on physics results is mitigated by the high granularity of the sub-detectors, allowing for some differentiation between particles from the primary interaction in a given bunch crossing and those from other interactions.

The Tracking System

The inner-most sub-detector of CMS is the inner tracking system. It is designed to precisely measure the trajectories of charged particles and reconstruct vertices from secondary decays. With a 5.8 m length, 2.6 m diameter, and about 200 m^2 of silicon area, it is the largest silicon tracker built to date. In order to reconstruct individual tracks, it has a high granularity and a fast response time. This is accomplished with three pixel layers and ten strip layers. To mitigate radiation damage from the high particle flux near the interaction point during Run 2, the entire tracker volume has been cooled to $-20°C$ and $-15°C$ for the pixel and strip layers respectively.

The extremely high particle flux near the interaction point requires the barrel pixel design to keep the occupancy at a manageable level. These first three pixel layers, with a radial distance of 2.2-10.2 cm from the interaction point, provide precise tracking points in r, ϕ, and z. This results in good impact parameter resolution as well as good secondary vertex reconstruction for b and τ decays. The pixel layers also seed tracks for reconstruction by the outer layers. Since the need to read-out all the channels as well as sheer cost make too many pixel layers unfeasible, the rest of the barrel is composed of 10 strip layers out to 1.1 m. Some of the modules are double-sided, with two modules set back-to-back but slightly offset, in order to measure both of the η-z coordinates. The tracker endcaps, which

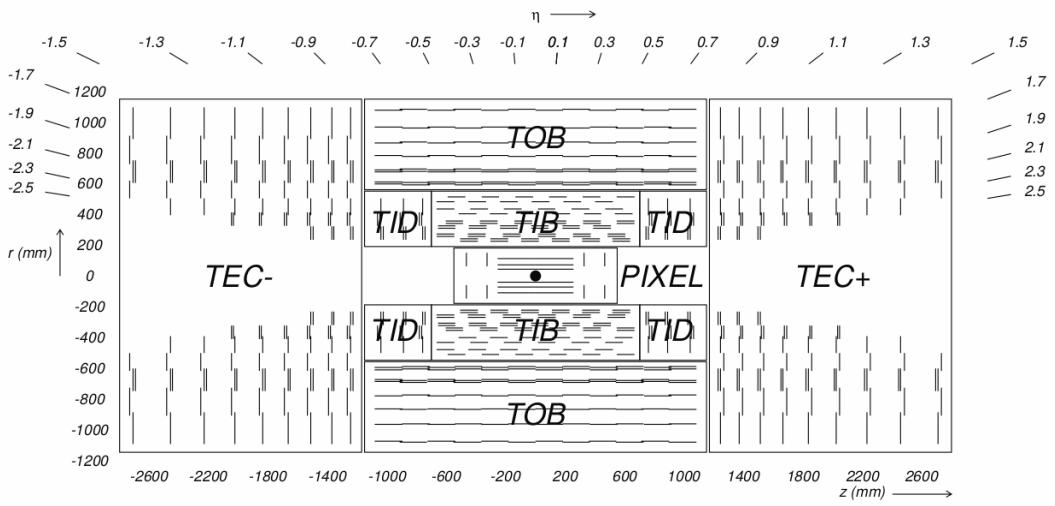

Figure 1.11: A schematic cross section of the tracker. Each module is represented by a line, with back-to-back modules represented by double lines. The dot at the center is the nominal interaction point. The pixel detector is the inner most box. The TIB/TID, TOB, TEC+, and TEC- are partitions of the strip detector.

are each composed of two pixel disks and 12 strip disks, extend the coverage to $|\eta| < 2.5$. A schematic cross section can be seen in figure 1.11.

The Electromagnetic Calorimeter

The next sub-detector is the electromagnetic calorimeter (ECAL). It is made up of over 75 thousand lead tungstate ($PbWO_4$) crystals as well as a preshower system installed in front of the endcaps. It covers the radial distance from 1.29-1.77 m and provides coverage up to $|\eta| < 3.0$. These high-density crystals are a good choice because they are radiation resistant, have a short radiation length, and about 80% of the light from incoming particles is emitted in the 25ns before the next bunch crossing. The crystals are 22-23 cm long, which is over 25 radiation lengths. Crystals in both the barrel and endcaps are mounted such that they are slightly offset from the vector pointing towards the nominal interaction point, as

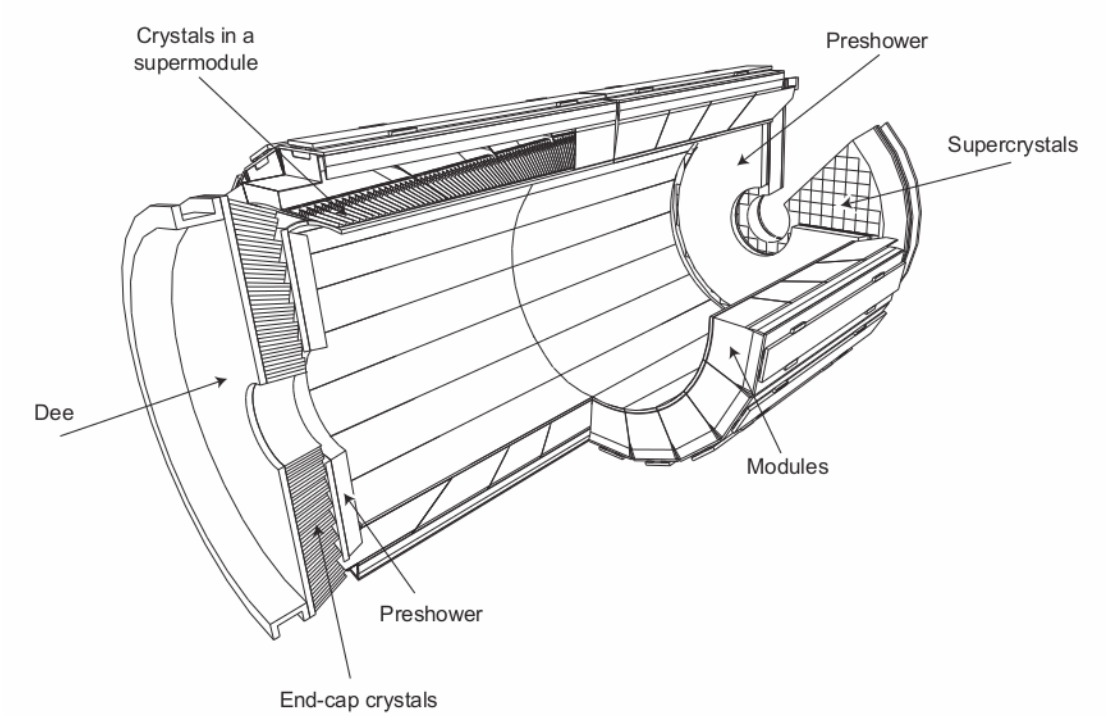

Figure 1.12: A cut-away drawing of the ECAL showing the crystal arrangement in the barrel and endcaps.

can be seen in figure 1.12, in order to avoid holes in coverage. Scintillation light is detected by silicon avalanche photodiodes and vacuum phototriodes in the barrel and endcaps respectively. This entire sub-detector is kept at a constant 18°C since the number of scintillation photons is temperature dependent.

The endcap preshower system is primarily to identify neutral pions, however, it also helps distinguish between electrons and minimum ionizing particles as well as improving the position resolution of electrons and photons. It is 20 cm thick and contains two layers that each consist of a lead radiator to initiate electromagnetic showers and silicon strips to measure the deposited energy and shower profiles.

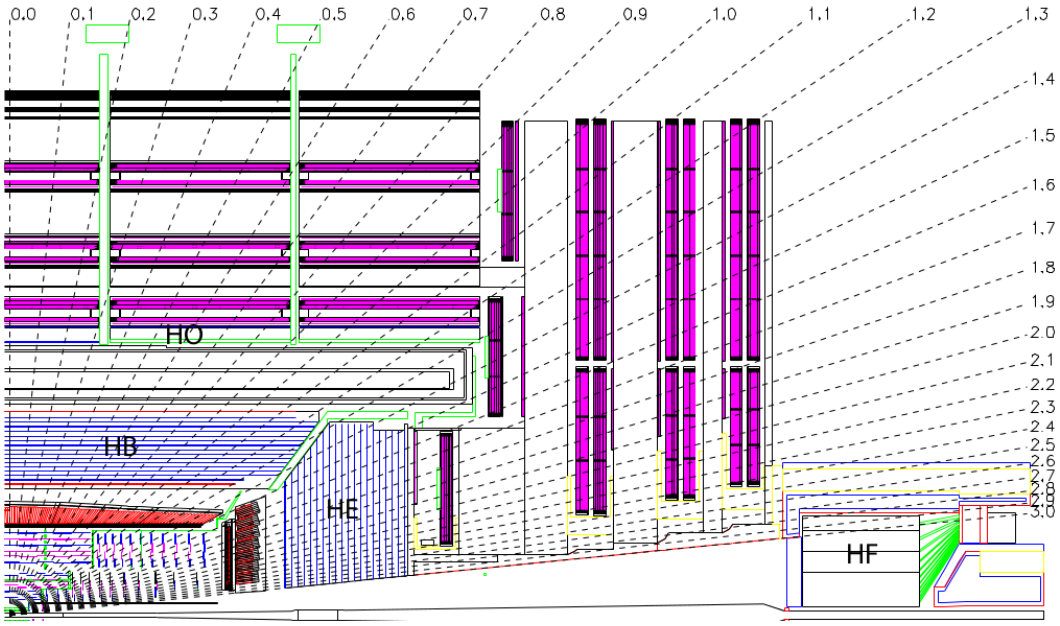

Figure 1.13: A cut-away drawing of one quarter of the CMS detector that shows the placement of the HCAL barrel (HE), endcap (HE), and tail catcher (HO) relative to the magnet and other sub-detectors.

The Hadronic Calorimeter

The last sub-detector before the solenoid, the hadronic calorimeter (HCAL), occupies the radial distances between 1.77-2.95 m. It is designed to measure the energy of hadronic jets. An outer tail catcher is also placed outside the magnet. This results in up to 11 and 9 hadronic interaction lengths in the barrel and endcap respectively, with the ECAL adding about one additional interaction length. The placement of these pieces can be seen in figure 1.13.

The barrel region, extending to $|\eta| < 1.3$, contains 72 azimuthal wedges. Each wedge consists of 14 brass absorber plates plus inner and outer stainless steel plates for structural strength. The plates are bolted together in a staggered configuration to eliminate dead areas and then aligned parallel to the beam line. Tiles of a moderately radiation hard plastic scintillator are used between the plates to

sample the hadronic showers. Since showers can start in the material between the ECAL and HCAL, scintillator tiles are also placed before the first steel plates. Wavelength shifting fibers embedded in the scintillator tiles are used to channel the light to photo detectors. The tail catcher is installed as the first detecting layer between the solenoid and the iron return yoke, taking advantage of the solenoid as up to an additional 1.4 interaction lengths.

The endcaps cover the region $1.3 < |\eta| < 3$ and use a similar scheme as the barrel with the geometry optimized to minimize gaps between the barrel and endcap. It also needs to be strong enough to support both itself and the ECAL endcap and preshower detector which are attached to the front. All of this must be accomplished with minimal structural material.

The Muon System

The final sub-detector is the muon system. Since muons are relatively easy to detect and lose less energy through radiation than electrons, they are useful for identifying interesting processes. Thus, as suggested by the experiment's name, the reconstruction of muons is extremely important. As the outermost layer, it has about 25,000 m^2 of detection planes. Thus it must be constructed from relatively inexpensive material. There are four stations integrated into the magnet's iron return yoke in the barrel region, as shown if figure 1.14. This takes advantage of the magnetic field in the return yoke to bend the muons' trajectories in order to get a precise momentum measurement. The iron yoke itself also acts as a hadron absorber to aid in muon identification. Three different types of detectors are used throughout the muon system.

Drift tube (DT) chambers with rectangular cells are used in the barrel ($|\eta| <$ 2.1) where the rate of muons is relatively low and the magnetic field is uniform.

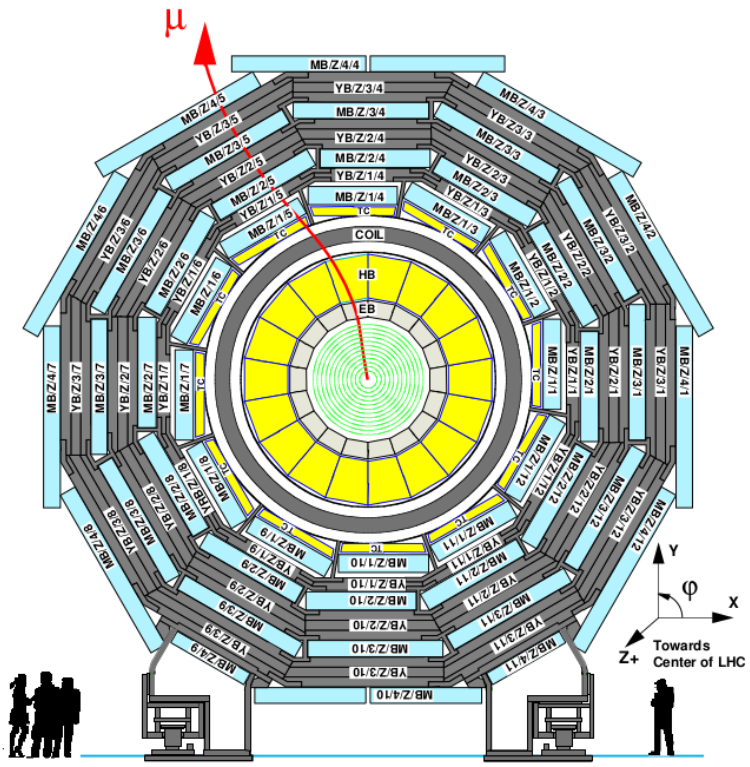

Figure 1.14: A cross section of the barrel of the CMS detector that shows how the muon system (light blue boxes, MB/Z/n/m) is integrated into the iron return yoke (dark gray boxes, YB/Z/n/m). From inside to outside, the sub-detectors are the tracker (green circles), ECAL (light gray boxes), HCAL (yellow boxes), and solenoid coil (dark gray disk). The red curve shows the path of a muon. through the detector.

Each of the first three stations contains two groups of four chambers which measure muon trajectories in the $r - \phi$ plane and 4 chambers for z measurements. The fourth station only measures r and ϕ. All cells are staggered by half a cell to eliminate dead spots. In the endcaps ($0.9 < |\eta| < 2.4$), both the muon and background rates are higher and the magnetic field is not uniform so cathode strip chambers (CSC) are used. CSCs are radiation resistant and have a fast response time and fine segmentation. Similar to the design of the barrel, the 4 CSC stations in each endcap are integrated between the solenoid's flux return

plates and, as with the other endcap detectors, in the plane perpendicular to the beam line. The cathode strips measure the r and ϕ coordinates while the anode wires measure η and the beam-crossing time.

The third muon detector, resistive plate chambers (RPC), is a dedicated triggering system with a fast response time and good time resolution. This comes at the cost of a coarser position resolution. One and two layers are embedded into each of the last and first two muon stations, respectively, in the barrel, and one layer in each of the first three stations in the endcaps. The RPCs are parallel-plate detectors with a two gaseous volumes, where the middle plate picks up and reads out the signal.

Triggering and Data Collection

With bunch crossings every 25 ns, there are on the order of 40 million events every second. Since it impossible to store and process anywhere near this amount of data, a reduction of about six orders of magnitude is necessary. This is accomplished with a two step process: the level-1 (L1) trigger and the high level trigger (HLT).

The L1 trigger uses programmable custom-designed electronics to reduce the rate to about 30 kHz. Coarsely segmented data from the calorimeters and muon system is used to make a decision on whether or not to keep an event in 3.2 μs while the full high-resolution data is stored in the front-end electronics. Due to the short timescale in which a decision must be made only minimal processing can be used, including no iterative algorithms. The final L1 trigger is built up through a series of steps that starts with local triggers. Also called trigger primitive generators, this is coarsely binned information about energy deposits in the calorimeters and track segments/hit patterns in the muon chambers that gets sent to the to the regional

triggers. Regional triggers combine the local information from their regions to sort and rank lists of trigger objects (e.g. electrons and jets). Each regional trigger sends the top candidates for each object to the global calorimeter and muon triggers as appropriate. The global triggers then determine the best global objects from each list as well as doing the final calculations for any global variables (e.g. total event energy). The final global trigger uses the information from the global calorimeter and muon triggers as well as information about the readiness of the sub-detectors in order to make the decision on whether or not to keep an event.

There can be up to 128 different L1 trigger algorithms. Each one combines various pieces of the information available to the global trigger using simple AND-OR logic in order to make the decision about keeping an event. Each one can also be independently prescaled to keep the final L1 rate within acceptable limits. Prescaling means only keeping a subset of the events that pass the trigger. For example, a prescale of 10 means that only every tenth event is kept. If any of the L1 algorithms pass (after prescaling) the event is kept.

The HLT uses a farm of commercial processors to reduce the rate of stored events to a few hundred per second. Although the HLT has access to the full event information, it does not have enough time to completely reconstruct every event and thus uses faster versions of the offline reconstruction. It uses the L1 trigger algorithm decision as seeds such that a given HLT algorithm only runs on events that pass its L1 seed. Since the HLT paths are analysis-specific, they will not be discussed further here.

1.3 Event Reconstruction

After an event has been read out of the detector and passed one or more of the HLT paths, the detector information needs to be reconstructed so that the event can be analyzed. The first step in this process is to use detector-level information such as tracker hits and calorimeter energy deposits to reconstruct particles. Once the individual particles have been constructed, they can be used to create jets, calculate lepton isolation, tag b-jets, and so on.

1.3.1 Particle Reconstruction

The particle reconstruction step is accomplished by using the particle-flow (PF) algorithm [20] to combine information from all the sub-detectors in order to reconstruct particles and determine their trajectory, type, and energy. This algorithm begins by reconstructing information from each sub-detector separately into tracks and calorimeter clusters which are then linked together to form blocks. The blocks are used to create the final list of particles in the event, including a preliminary determination of their type (muon, electron, photon, charged hadron, neutral hadron) and energy.

The ability to reconstruct charged particle tracks with a high efficiency and low fake rate makes the tracker the ideal place to begin reconstructing an event. This is done by using an iterative process that creates the highest quality tracks possible and then removes all tracker hits associated with the newly created tracks before repeating the process. In subsequent iterations, the criteria for what constitutes a good track is loosened in order to increase the overall reconstruction efficiency. This does not drastically increase the fake rate since the combinatorics have been reduced by the removal of all hits used in the previous iterations. The last two

iterations relax the constraints on how close the vertex of a track is required to be to the interaction point in order to also reconstruct charged particles from secondary decays.

Next, energy deposits in the calorimeters are combined into clusters. The same process is implemented separately for the ECAL and HCAL sub-detectors. First, local energy maxima are used to identify cluster seeds. The seeds are grown into clusters by iteratively adding adjacent cells until none with an energy sufficiently above the noise threshold remain. If more than one seed is in a given cluster the energy is iteratively divided between the seeds based on relative distance.

Once all the local sub-detector objects are created they are linked together. Clusters are linked between calorimeters when they are sufficiently overlapping. Tracks are extrapolated from the last measured hits into the calorimeters and linked to any clusters they intersect. To account for Bremsstrahlung radiation, tangents to the tracks are also extrapolated to the ECAL and linked with any intersecting clusters. Tracks in the tracker and muon system are linked when the fit between the two is sufficiently good based on a χ^2 test.

Each block is then analyzed separately to create the final particle list. For a given block if there are any links between tracker and muon tracks the compatibility of the momentum calculated using only the tracker track is compared to the momentum calculated using both linked tracks. If they are compatible both tracks are removed from the block and a particle-flow muons (PF muon) is created. Next, an attempt is made to extrapolate any tracks that are likely to be electrons (shorter tracks with energy lost to Bremsstrahlung radiation) to an ECAL cluster. If a final particle-flow electron (PF electron) identification is made, the corresponding track and ECAL clusters are removed from the block. All remaining tracks are required to pass a tighter set of quality criteria to reduce

fakes before additional links between tracks and calorimeter clusters are analyzed. For each track, any associated ECAL and HCAL clusters up to the track's momentum are identified as particle-flow charged hadrons (PF charged hadron) and removed from the block. All remaining ECAL and HCAL clusters are identified as particle-flow photons (PF photon) and particle-flow neutral hadrons (PF neutral hadron) respectively. All the above particle-flow particles are collectively referred to as PF candidates.

1.3.2 Jet Clustering

After all the individual particles are reconstructed, they are clustered together into jets to reconstruct the hadronization of quarks and gluons. This is done by using the anti-k_t clustering algorithm [21]. This algorithm begins by creating a list of "distances" containing the distance between between each pair of particles:

$$d_{ij} = \min\left(k_{t,i}^{2p}, k_{t,j}^{2p}\right) \frac{(\Delta R)^2}{R^2}$$

and between each particle and the beam:

$$d_{iB} = k_{t,i}^{2p}$$

where $k_{t,i}$ is the transverse momentum of particle i. The distance parameter R is a measure of how much area jets tend to encompass and is currently set to 0.4 for standard jets. p is used to determine the relative influence of the energy versus geometrical scales as well as to decide whether high or low momentum jets are prioritized by the algorithm to cluster first. The anti-k_t algorithm sets $p = -1$ so that higher momentum jets are clustered first.

Each itteration then begins by finding the smallest distance on the current list. If it is a d_{ij}, particles i and j are combined and all distances involving them are recalculated using the new combined object. Otherwise, the particle associated to d_{iB} is designated as a jet and all associated distances are removed from the list. This process is repeated until no distances are left.

1.3.3 b-Jet Identification

The relatively long lifetime of b hadrons means that jets that originate from b-quarks are displaced from the primary interaction vertex, instead creating a secondary vertex. The Combined Secondary Vertex (CSV) tagger [22, 23] takes advantage of this feature in order to identify jets initiated by b-quarks. This algorithm has been re-optimized for Run 2 (CSVv2). It combines information about track impact parameters and the secondary vertices associated to a jet using a multivariate technique. The output of the CSVv2 algorithm is a discriminant which is a measure of how likely it is that a given jet originated from a b-quark. Several different values of the discriminant, which are called working points, are recommended for use in analyses depending on the specific analysis's needs for a pure sample of b-jets versus a high efficiency of collecting any b-jets. These working points are labeled as loose, medium, and tight in order of increasing b-jet purity.

1.3.4 Top-Tagging

The CMS Top Tagging (CTT) algorithm [24,25] is used to reconstruct boosted hadronically decaying top-quarks, i.e. when the W produced in the initial decay subsequently decays into two quarks. In order to collect these three decay products

into a single jet "hard jets" are defined as jets that are clustered with a distance parameter that is twice as large as that of standard jets ($R = 0.8$). These hard jets are then broken up into subjets using both the distance between and the p_{T} fraction of subjet candidates. If two subjets are successfully found, an attempt is made to further break each of them up into subjets as well. Thus, depending on whether or not each subjet is successfully broken up into subjets, the original hard jet can have up to four reconstructed subjets.

In order for a given hard jet to be tagged as a top, it must have a reconstructed mass near the top mass and have at least three subjets with the minimum pairwise invariant mass near the W mass. Using top-tagged hard jets with low p_{T} ($\lesssim$ 350 GeV) is not recommended since the low momentum decay products of such a top are unlikely to be merged into a single hard jet.

1.3.5 Electron Identification

Additional criteria are applied to PF electrons as simple cuts on observables in order to identify them as electrons that are good for an analysis. The cut thresholds increase in strictness in order to identify them as, in order of increasing purity, veto, loose, medium, and tight electrons. The observables used include: a variable that probes the width of the ECAL cluster in the η direction, how well the track and ECAL cluster are matched in ϕ and η, the ratio of energy deposited in the HCAL behind the electron's ECAL cluster over the energy deposited in the ECAL cluster, the distances in the transverse plane and along the beam axis between the beginning of the electron's track and the primary vertex, and the maximum number of allowed missing tracker hits.

1.3.6 Muon Identification

Muon identification begins with PF muons as the loosest identification requirement. Medium and tight identifications are also defined with stricter cuts on the quality of the track fit and number of hits in the various tracker and muon layers.

1.3.7 Photon Identification

Additional criteria are applied to PF photons as simple cuts on observables in order to identify them as photons that are good for an analysis. The cut thresholds increase in strictness in order to identify them as, in order of increasing purity, loose, medium, and tight photons. The observables used include: a variable that probes the width of the ECAL cluster in the η direction, the ratio of energy deposited in the HCAL behind the electron's ECAL cluster over the energy deposited in the ECAL cluster, and the isolation of the photon with respect to charged hadrons, neutral hadrons, and other photons.

Chapter 2

Search Design

Although 3.8 fb^{-1} of data has been recorded by CMS, ongoing technical difficulties resulted in the magnet being off for significant portions of the data-taking period. Thus only 2.3 fb^{-1} of the full 2015 dataset are useful for most analyses including this one. To analyze this data for events with stop decays, a baseline search region is defined in section 2.3 with a set of cuts on standard objects and variables. The baseline selection is designed to create a region where the fairly small number of potential signal events are not completely hidden by SM processes that have cross sections many orders of magnitude higher than signal processes. This selection is based on the most distinct characteristics of signal events: high $\not{E}_{\text{T}}$ and the number of jets and b-jets in an event. Section 2.1 gives an overview of the data and simulated samples used in this analysis. Any analysis specific decisions about the reconstructed objects used in this analysis are in section 2.2.

In section 2.4 the baseline search region is divided into a set of bins using both tighter selections on some of the observables in the baseline selection and additional observables. Although these tighter selections can provide substantial discriminating power for some signal models, for other models simply adding

tighter cuts would also exclude a substantial portion of potential signal events. Binning rather than tightening the search region both increases the statistical power of the search and limits the extent to which signal events are lost.

After the baseline selection, there are still considerable contributions from several processes in the search region. The most significant contribution is from events with $t\bar{t}$ decays and, to a lesser extent, events with a W decay plus additional jets. These events primarily enter the search region when a W boson decays leptonically but the lepton is not identified as such. This background is discussed further in section 3.1. The next most predominant background, discussed in section 3.2, comes from events where a Z boson decays into two neutrinos. When this happens in an event that also has a significant number of jets the $\not{E}_{\mathrm{T}}$ produced by the neutrinos can cause the event to look similar to signal events. Although the baseline selection almost completely eliminates QCD multijet events, this source of background is not trivial. The jet mis-measurements that result in enough $\not{E}_{\mathrm{T}}$ for events to enter the search region are difficult to model, as discussed in section 3.3. The final process with noticeable contributions to the search region is ttZ. These events have both two fully hadronic $t\bar{t}$ decays and an invisibly decaying Z boson. Although it is completely irreducible due to the real top decays and real $\not{E}_{\mathrm{T}}$, its cross-section is sufficiently low that it is negligible in all but the most exclusive bins. Section 3.4 further discusses this background.

The contributions from each of these background processes to each bin of the search region are estimated separately using the dedicated methods described in chapter 3. Each method follows a similar procedure where a control region that is rich in the relevant processes is first defined. Simulated events are then used to make a prediction for the number of events expected in the search region based on the number of events observed in the control region. Since the simulation

of various background processes is relied on for this extrapolation, ensuring that the data is correctly modeled is an important aspect of the analysis. This is accomplished by reweighting the simulated events with various correction factors so that they match the data better as discussed in section 3.5. Any uncertainties on these corrections as well as other systematic effects are estimated by varying relevant parameters by their uncertainties and propagating the variation to the final background prediction. The percent change in the prediction from a variation is then taken as a systematic uncertainty on the prediction.

In order to avoid introducing any biases into the search, the data in the baseline search region is blinded while the search is designed and optimized. The full background estimation methods are tested in regions that are both adjacent to and distinct from the search region, as discussed in section 3.6. After every piece of the analysis is in place it is presented to the CMS SUSY group for an initial internal review. After approval is given to unblinded the search region, the observed number of each events in every bin is compared to the predicted number. These results are given in chapter 4. Finally, in section 4.1, the results are interpreted in the context of the simplified models that were defined in section 1.1.3.

2.1 Data and Simulated Samples

2.1.1 Datasets

The distinguishing characteristics of stop pair production are missing energy ($\not\!\!E_{\mathrm{T}}$, fully defined in section 2.2.3) from the LSP and multiple jets from top decays. Therefore, $\not\!\!E_{\mathrm{T}}$ and jets are good observables to use to collect the primary dataset for this analysis. At the L1 trigger stage all events with $\not\!\!E_{\mathrm{T}} > 60$ GeV are kept.

The choice of 60 GeV for the $\not\!\!E_{\mathrm{T}}$ threshold keeps the rate of this trigger within the allocated amount. This seeds an HLT path that uses both $\not\!\!E_{\mathrm{T}}$ and jets, requiring at least two jets with $p_{\mathrm{T}} > 55$ GeV and $\not\!\!E_{\mathrm{T}} > 110$ GeV.

Some events which would have $\not\!\!E_{\mathrm{T}} > 110$ GeV and at least two jets with $p_{\mathrm{T}} > 55$ GeV after being fully reconstructed will not meet these thresholds with the courser calculations used in the trigger algorithms. The inefficiency in collecting events is measured using an independent dataset which is collected by triggering on a single electron. The number of events in the single electron dataset that also pass the $\not\!\!E_{\mathrm{T}}$+jets trigger is looked at as a function of the triggering variables, $\not\!\!E_{\mathrm{T}}$ and the p_{T} of the two highest p_{T} jets. Almost all events pass the $\not\!\!E_{\mathrm{T}}$+jets trigger after requiring $\not\!\!E_{\mathrm{T}} > 250$ and two jets both with $p_{\mathrm{T}} > 75$. The primary source of the inefficiency for lower $\not\!\!E_{\mathrm{T}}$ events to pass the trigger is the inefficiency of the L1 trigger due to the coarseness of the information available to it. In the $200 \leq \not\!\!E_{\mathrm{T}} < 250$ region, which is used as a validation region, over 95% of events pass the trigger. The trigger efficiency as a both function of $\not\!\!E_{\mathrm{T}}$ and the second jet's p_{T} is shown in figure 2.1.

Control regions collected with other triggers are used for various studies and background estimations. Datasets for leptonic control regions are collected with single muon, single electron, double muon, double electron, and muon plus electron triggers. In particular, the double electron and double muon triggers are used to obtain the $\mathrm{Z} \to \ell\ell$ sample used for the $\mathrm{Z} \to \ell\ell$ prediction. The single photon dataset used for the $\mathrm{Z} \to \ell\ell$ prediction is collected with a single photon trigger. Several datasets collected on H_{T} triggers with varying thresholds are used for studies related to the QCD prediction. Due to the increased rate of these triggers as the H_{T} threshold drops, the lower threshold triggers are prescaled. Therefore these datasets need to be patched together to obtain a continuous H_{T} distribution.

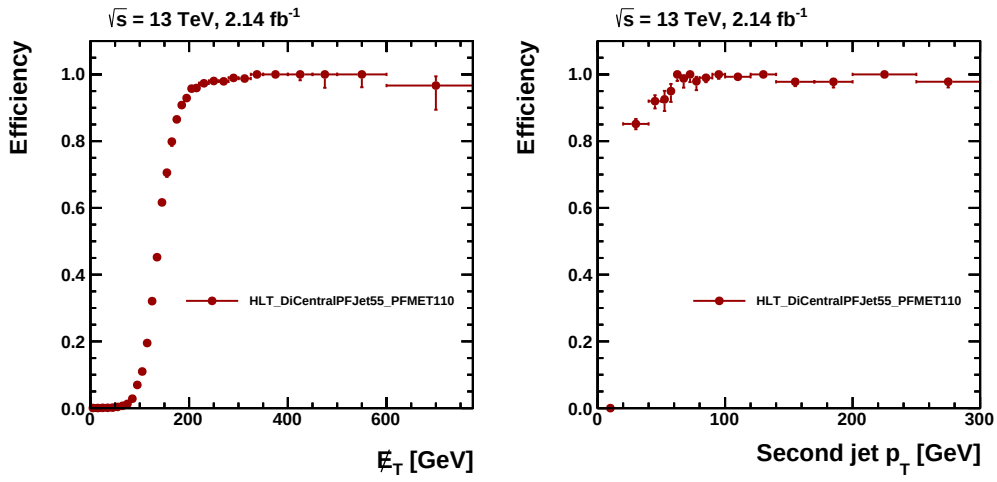

Figure 2.1: Trigger efficiency as as a function of $\not{E}_{\mathrm{T}}$ after requiring two jets with $p_{\mathrm{T}} > 75$ GeV (left) and as a function of the second jet's p_{T} after requiring $\not{E}_{\mathrm{T}} > 250$ (right).

Several filters are applied to reject events that have unphysical observables such as 20 TeV of $\not{E}_{\mathrm{T}}$ from various sources which include beam halo interactions, electronic noise, and bad detector regions.

2.1.2 Simulated Events

Monte Carlo (MC) simulated events are a useful tool for understanding various processes. They can be used to design a search well before any data is available. Once data is available, they can also be used to make predictions about what the data in a given search region will look like based on observations of data in control regions. The MC events used in this analysis are generated centrally by a dedicated group using a multi-step process. Datasets are simulated for a wide variety of Standard Model and non-Standard Model processes. The final files can be analyzed as if they were actual data files.

First, matrix-element generators including MadGraph5 [26], Powheg [27–

30], and MadGraph5_AMC@NLO [31] use matrix element calculations to generate the primary parton interactions in each event. These events are then passed to a general purpose generator called Pythia8.1 [32] for parton showering and hadronization. Additional parton interactions are added to account for spectator interactions of the partons not involved in the hard scatter. Pileup interactions are also added by simulating additional inelastic collisions. This is usually done before data-taking and thus before the distribution of the number of pileup interactions in data is known. To account for this MC samples will be reweighted at the analysis level so that their pileup distribution matches data. The final step is to simulate the response of the CMS detector response with software based on Geant4 [33] to get the final set of events.

Various corrections are applied to MC samples at the analysis level to obtain better agreement with data. These are discussed in the relevant sections of chapter 3, particularly section 3.5, for background samples and in section 4.1 for signal samples.

Signal samples are generated in a grid to cover the $m_{\tilde{t}} - m_{\tilde{\chi}_1^0}$ plane in 50 GeV intervals. Near the $m_{\tilde{t}} = m_{\mathrm{t}}$ diagonal, the binning is increased to 25 GeV intervals. For T2tb samples, a 50% branching ratio is used for both T2tt and T2bW, and the latter is generated assuming $m_{\tilde{\chi}^{\pm}} - m_{\tilde{\chi}_1^0} = 5$ GeV. In order to have a reasonable amount of statistics in every bin of this grid without requiring massive amounts of computing power, a lighter version of the simulation software (FastSim) [34] is used. Additional corrections are derived to correct FastSim samples to fully simulated samples using samples generated at a few representative signal points.

2.2 Object Definitions

This section describes object selection requirements that are not part of the standard event reconstruction in section 1.3. It also describes any additional selection requirements on and specific choices for the standard objects. Most standard objects have multiple sets of identification criteria for different levels of sample purity. This is to cater to different needs for minimizing the number of object candidates that are not identified verses ensuring that the selected object candidates have a low number of fakes, i.e. efficiency verses purity. The different sets of selection criteria are provided centrally for CMS by groups of experts in the various objects. Corrections needed to adjust MC to better match data related to these objects are also mentioned.

2.2.1 Vertices

Vertices are constructed by fitting the trajectories of particles using information from the tracker, especially the pixel layers. Events are only considered if they have at least one reconstructed vertex that is within 24 cm of the center of the detector along the z-axis and within 2 cm of the beamline in the transverse plane. The vertex with the highest Σp_{T}^2 of tracks associated to it that passes the the previous distance requirements is designated as the primary vertex.

2.2.2 Jets

On top of the standard anti-k_t jet clustering described in section 1.3.4, a charged hadron subtraction procedure is used to correct jets for pileup (PU). This is done by removing charged hadrons that can be associated to reconstructed vertices other than the primary vertex. Centrally supplied jet energy corrections

are also applied to correct for any remaining PU effects as well as variations in the detector response and differences between data and MC.

Only jets with $p_{\mathrm{T}} > 20$ GeV that are within the tracker ($|\eta| < 2.4$) and satisfy a loose set of centrally supplied jet identification criteria are considered. Keeping jets with this low p_{T} threshold results in better efficiency for signal models with lower mass splittings that can have soft decay products.

2.2.3 Missing Transverse Energy

$\not{E}_{\mathrm{T}}$ is defined as the negative vectorial sum of the transverse momentum of all PF candidates. Centrally provided corrections based on the standard jet energy corrections are applied to the raw PF $\not{E}_{\mathrm{T}}$.

2.2.4 b-Tagging

Both loose and medium working points of the CSVv2 discriminator are used to identify b-jets, as mentioned in section 1.3.3. Centrally supplied scale factors are applied to MC to correct for efficiency differences between data and MC.

2.2.5 Top-Tagging

Tagged tops, as defined in section 1.3.4, are used if the hard jet's reconstructed mass is within the window 140-250 GeV and it has at least three subjets with the invariant mass of every pair of the three highest p_{T} ones above 50 GeV. They are also required to have $p_{\mathrm{T}} > 400$ GeV, and be within the tracker ($|\eta| < 2.4$). Corrections to match MC to data are discussed in section 3.5.5 and applied as necessary.

2.2.6 Electrons and Muons

Electrons and muons, as discussed in sections 1.3.5 and 1.3.6 respectively, are required to have $p_{\mathrm{T}} > 5$ GeV, $|\eta| < 2.4$ and are identified using different selection criteria for the search and control regions. A loose set of veto criteria is used to exclude any events with a potential electron or muon candidate from the search region. Much tighter criteria are used to ensure a high purity sample of leptons for leptonic control regions.

Since this analysis is primarily interested in leptons that are not from heavy flavor decays (prompt leptons), all muon candidates are required to originate from a point near the primary vertex. Similar requirements for electrons are already included in the standard identification requirements.

To further aid in the identification of prompt leptons, all lepton candidates are required to be isolated from other activity in the event. The quantity used to determine a lepton's local isolation, called mini-isolation, is computed as the scalar sum of the p_{T} of all PF candidates within a ΔR cone of the lepton candidate. The size of the cone decreases as the lepton's p_{T} increases in order to contain the products of leptonic b-decays while still reducing overlaps with jets in the event. This is calculated after subtracting off the expected contribution from pileup. Electron and muons are considered to be isolated if their mini-isolation relative to the lepton p_{T} is less than 0.1 and 0.2 respectively in both the search and control regions.

2.2.7 Taus

A dedicated custom veto is used to identify and then reject events with hadronically decaying τ leptons. Hadronically decaying τ candidates are PF charged

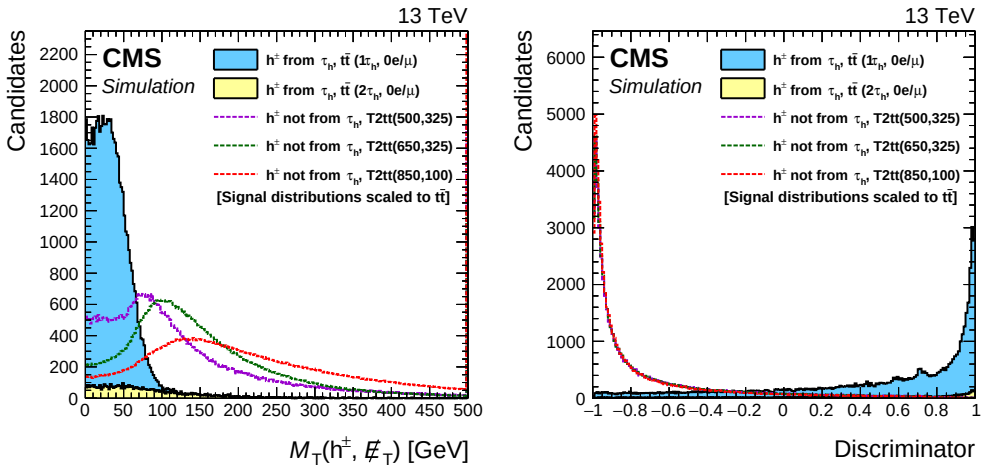

Figure 2.2: Distribution of $M_{\mathrm{T}}\left(\tau, \not{E}_{\mathrm{T}}\right)$ (left) and the τ veto BDT discriminant distribution (right) for charged hadrons. Candidates from τ decays in $\mathrm{t\bar{t}}$ MC are solid histograms while all non-τ charged hadrons from three example all-hadronic signal samples are lines. A preselection of $\not{E}_{\mathrm{T}} > 150$ GeV and $N_{\mathrm{J}}(p_{\mathrm{T}} > 30) \geq 4$ is applied. Each signal distribution is scaled to the area of the total $\mathrm{t\bar{t}}$ distribution.

hadrons with $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.4$, and an impact parameter along the z-axis $|d_z| < 0.2$ cm. Since τ decays often include neutral pions it is helpful to include the highest p_{T} photon within a small cone of the charged hadron in the τ candidate's four-momentum. Since a real hadronic τ decay will have an endpoint at the mass of the W boson, the transverse mass can be used to identify hadronic tau decays. $\not{E}_{\mathrm{T}}$ is used in place of the undetectable neutrino in the calculation. Thus $M_{\mathrm{T}}\left(\tau, \not{E}_{\mathrm{T}}\right)$ is defined as:

$$M_{\mathrm{T}}\left(\tau, \not{E}_{\mathrm{T}}\right) = \sqrt{2 \cdot p_{\mathrm{T}}\left(\tau + \text{nearest } \gamma\right) \cdot \not{E}_{\mathrm{T}} \cdot \left(1 - \cos\Delta\phi\right)}$$

and shown in the left plot of figure 2.2 for the $\mathrm{t\bar{t}}$ background and several example signal points. All τ candidates are required to have $M_{\mathrm{T}}\left(\tau, \not{E}_{\mathrm{T}}\right) < 100$ GeV.

After applying the above selection requirements to τ candidates, the final

selection is made using a multi-variate Boosted Decision Tree (BDT) [35]. A BDT combines the separation power of several variables that discriminate between charged hadrons from truth-matched τ decays in $t\bar{t}$ events and signal events without hadronic τ decays.

In a Decision Tree, a binary sequence of splits is used to separate signal from background. All events start in the first node where the best discriminating variable is used to split all events into two new nodes. Each new node uses the best discriminating variable for its subset of events to split them between two additional notes. This process is continued until each node is either all signal, all background, has too few events to split, or a maximum depth of splits is reached. Every final node is then classified as either signal or background based on the majority of events in it. The same tree of cuts can then be applied to a new set of events to obtain a classification of either signal or background for each event.

In a Boosted Decision tree, after a given tree is trained, all events that were misclassified by the tree have their weights increased and a new tree is trained. This is repeated many times and a final continuous discriminator in the range $[-1, 1]$ is obtained based on how often an event is categorized as signal or background.

A pre-selection with $\not{E}_{\mathrm{T}} > 150$ GeV and at least four jets with $p_{\mathrm{T}} > 30$ GeV is required for the BDT training. This is to partially replicate the environment of the search region (which will be fully described in section 2.3). The following variables are used for the τ veto BDT:

- The τ candidate's p_{T} and $|\eta|$.
- The scalar sum of the p_{T} of all charged particles that are both associated to the primary vertex and within ΔR cones around the τ candidate with sizes

0.1, 0.2, 0.3, and 0.4.

- The scalar sum of the p_{T} of all particles within the same ΔR cones as above of the τ candidate after correcting for PU.
- The distance in ΔR between the τ candidate and the nearest charged PF candidate with $p_{\mathrm{T}} > 1$ GeV.
- The distance in ΔR between the τ candidate the jet containing it. If the jet has $p_{\mathrm{T}} > 30$ GeV, the b-tagging discriminant for the jet is also included.

The τ veto BDT discriminant is shown on the right side of figure 2.2. All PF charged hadrons with $p_{\mathrm{T}} > 10$ GeV, $|\eta| < 2.4$, $|d_z| < 0.2$ cm, $M_{\mathrm{T}}\left(\tau, \not{E}_{\mathrm{T}}\right) < 100$ GeV, and a BDT discriminant greater than 0.55 are considered to be isolated hadronic taus.

2.2.8 Photons

Photons are required to pass a loose selection criteria as discussed in section 1.3.7 and have $p_{\mathrm{T}} > 180$ GeV and $|\eta| < 2.5$. The high p_{T} requirement is dictated by the high p_{T} threshold of the single photon trigger.

2.3 Baseline Selection

The signals targeted by this search occur at rates that are orders of magnitude lower than many SM processes. In order to be sensitive to signal events a set of cuts called the baseline selection is designed to remove the majority of background events while maintaining a high efficiency for signal events. As discussed at the end of section 1.1.3, a stop is expected to either decay directly to a top and LSP

	$t\bar{t}$/W+jets		$Z \to \nu\nu$		QCD		ttZ		T2tt(700,100)		T2tb(700,100)		T2tt(250,150)	
Trig. eff.	21961		7161		18376		71		102		102		1222	
e/μ veto	9015	41%	7090	99%	17750	97%	39	55%	64	63%	72	71%	644	53%
τ veto	6566	73%	6975	98%	16670	94%	35	89%	61	96%	67	94%	541	84%
$N_J \geq 5$	2895	44%	1722	25%	5544	33%	28	80%	54	87%	49	72%	459	85%
$N_b^l \geq 2$,	1470	51%	390	23%	2066	37%	21	75%	43	80%	38	79%	204	44%
$N_b \geq 1$	1195	81%	198	51%	1161	56%	18.8	90%	39	93%	35	92%	148	73%
$\Delta\phi_{1234} > 0.5$	580	49%	138	70%	32	3%	13.5	72%	33	84%	28	79%	56	38%

Table 2.1: The effect of the baseline selection cuts on each of the four main backgrounds and a few example signal points. Each efficiency is with respect to the previous row. "Trig. eff." refers to the cuts on $\not{E}_T$ and N_J to ensure the trigger is almost fully efficient ($\not{E}_T > 250$, $p_T(J_{1,2}) > 75$).

($\widetilde{t} \to t\widetilde{\chi}_1^0 \to bW^{\pm}\widetilde{\chi}_1^0$) or through an intermediate chargino ($\widetilde{t} \to b\widetilde{\chi}^{\pm} \to bW^{\pm}\widetilde{\chi}_1^0$). This search is only concerned with the fully hadronic channel therefore only events in which both of the W bosons decay into quarks are of interest.

The final all-hadronic observable state for events in which two stops are produced has two b-jets, four other jets, and considerable missing momentum. This makes observables such as $\not{E}_T$ and the number of jets/b-jets in an event powerful discriminators between signal and background events. Thus they are a good choice for the foundation of the baseline selection. This section describes every cut in the baseline selection as well as which types of background events each is intended to reduce. Table 2.1 lists all the baseline selection cuts and shows their effect on the primary backgrounds and three example signal points.

The first cuts in the baseline selection, $\not{E}_T > 250$, $p_T(\text{J1}) > 75$ GeV, and $p_T(\text{J2}) > 75$ GeV, are to maintain a high efficiency for passing the trigger as discussed in section 2.1.1. J1 and J2 refer to the jets with the highest and second highest p_T respectively. The $\not{E}_T$ distribution is shown in figure 2.3 and the distribution of N_J for jets with $p_T > 75$ GeV is shown in the left plot of figure 2.4.

As this search only targets fully hadronic decays, all events are required to

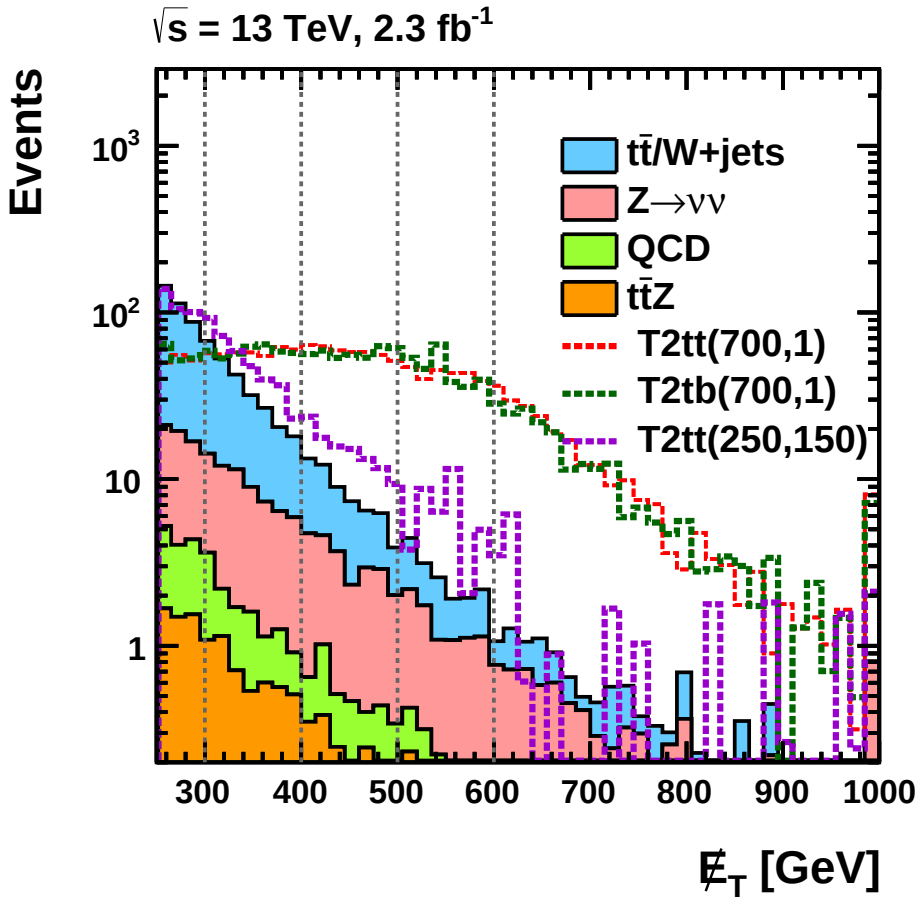

Figure 2.3: $\not{E}_{\mathrm{T}}$ after applying the full baseline selection for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The signals are each normalized to the total background and the gray lines show the binning.

have no isolated electrons, muons, or taus as defined in sections 2.2.6 and 2.2.7. The e/μ and τ vetoes also reduce the background from $\mathrm{t\bar{t}}$ and W+jets events by about half.

Since both the targeted models have six jets in the final state at least five jets are required in each event. A sixth jet is not required in order to maintain signal efficiency in the case where one jet is not reconstructed because it is soft ($p_{\mathrm{T}} < 20$) or out of acceptance ($|\eta| > 2.4$). All backgrounds except ttZ are reduced by over

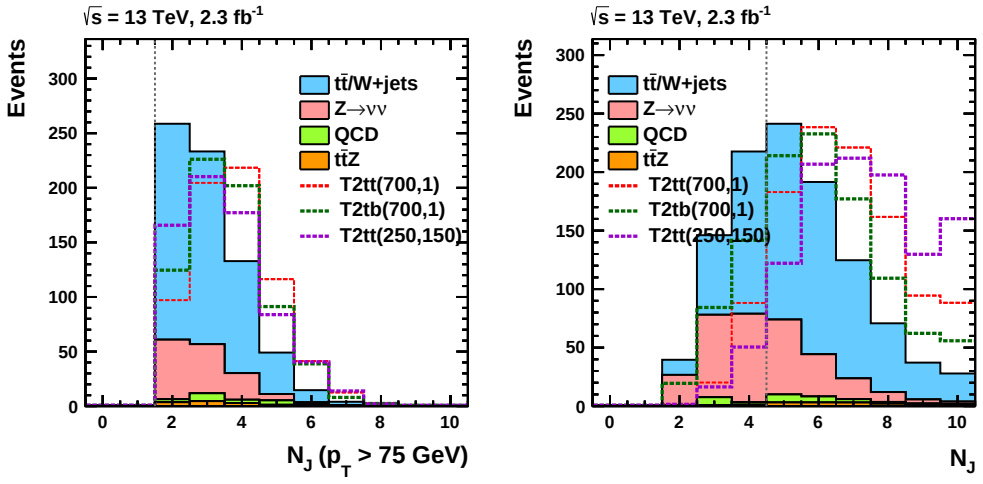

Figure 2.4: Distributions of $N_{\mathrm{J}}(p_{\mathrm{T}} > 75$ GeV) (left) and $N_{\mathrm{J}}(p_{\mathrm{T}} > 20$ GeV) (right) for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The trigger efficiency cuts ($\not\!\!E_{\mathrm{T}} > 250$, $N_{\mathrm{j}}(p_{\mathrm{T}} > 75) \geq 2$) and all the baseline cuts except the one for the shown observable are applied. The signals are each normalized to the total background and the gray lines show the selection threshold for each observable.

a factor of two. The distribution of N_{J} (for jets with $p_{\mathrm{T}} > 20$ GeV) is shown in the right plot of figure 2.4.

Both decay models also produce two b-quarks, either directly or through the decay of a top quark. Therefore, all events are required to have at least two of their jets loosely identified as b-jets: $N_{\mathrm{b}}^{l} \geq 2$ where N_{b}^{l} is the number of jets that pass the loose working point of the CSVv2 discriminator. This drastically reduces all backgrounds except ttZ. The backgrounds are further reduced with only a small effect on potential signal events by requiring one of the b-jets to pass a stricter identification: $N_{\mathrm{b}} \geq 1$, where N_{b} is the number of jets that pass the medium working point of the CSVv2 discriminator. The distributions of the number of loose and medium b-jets are shown in figure 2.5.

One source of fake $\not\!\!E_{\mathrm{T}}$ is from a substantial amount of a jet's energy being missed by the detector. This results in potentially substantial $\not\!\!E_{\mathrm{T}}$ values that are

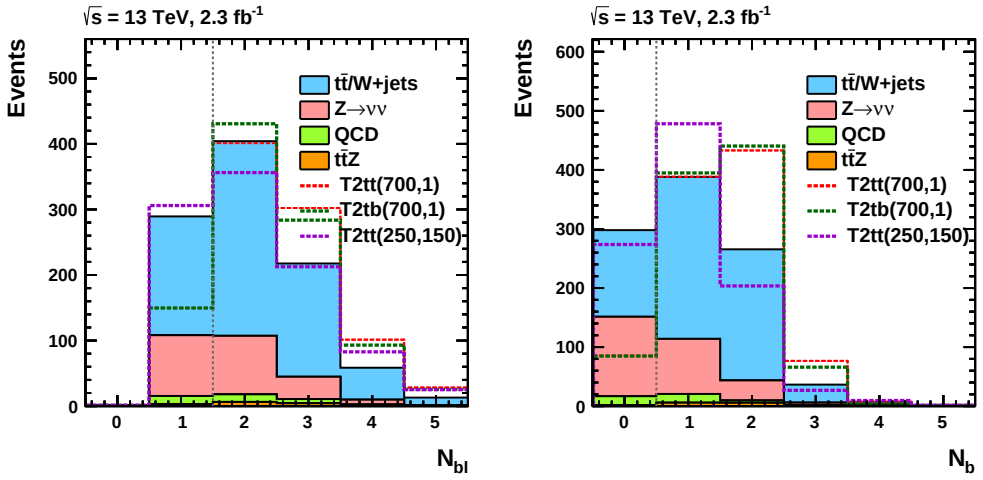

Figure 2.5: Distributions of N_{b}^{l} (left) and N_{b} (right) for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The trigger efficiency cuts ($\not{E}_{\mathrm{T}} > 250$, $N_{\mathrm{j}}(p_{\mathrm{T}} > 75) \geq 2$) and all the baseline cuts except the one for the shown observable are applied. The signals are each normalized to the total background and the gray lines show the selection threshold for each observable.

along nearly the same vector as the jet. To protect the search region from these events, the final cuts of the baseline selection require the azimuthal angles between $\not{E}_{\mathrm{T}}$ and each of the four highest p_{T} jets in every event ($\Delta\phi_{1234}$) to be above 0.5. This drastically reduces the QCD multijet background which has no real sources of $\not{E}_{\mathrm{T}}$ (e.g. from neutrinos). The $\Delta\phi$ distributions are shown in figure 2.6.

2.4 Event Categorization

Tightening the cuts on some of the observables used in the baseline selection as well as introducing requirements on additional observables can provide significantly more discriminating power for some signal points. However, simply adding and tightening simple cuts would also drastically reduce the efficiency for other signal points, particularity those with lower mass differences between the stop and

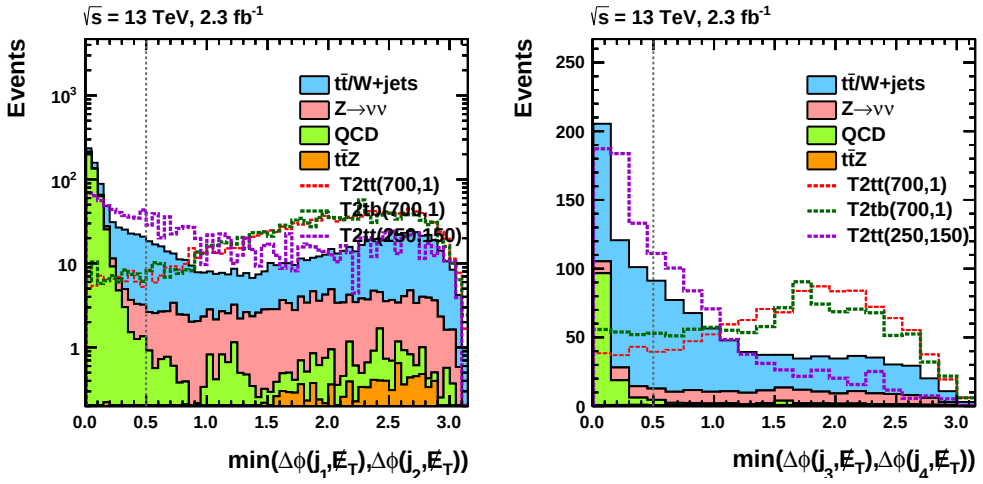

Figure 2.6: Distributions of the minimum of $\Delta\phi$ between the two highest p_{T} jets (left) and third/fourth highest p_{T} jets (right) for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The trigger efficiency cuts ($\not{E}_{\mathrm{T}} > 250$, $N_{\mathrm{j}}(p_{\mathrm{T}} > 75) \geq 2$) and all the baseline cuts except the one for the shown observable are applied. The signals are each normalized to the total background and the gray lines show the selection threshold for each observable.

LSP. In order to benefit from the additional separation power without any loss in efficiency, the baseline search region is divided into a series of bins as follows.

As the most discriminating observable for signal points with large mass splittings, the $\not{E}_{\mathrm{T}} > 250$ GeV region is divided into five distinct regions: $[250, 300)$, $[300, 400)$, $[400, 500)$, $[500, 600)$, ≥ 600 GeV. The $\not{E}_{\mathrm{T}}$ distribution after the baseline selection for the backgrounds and a few example signal points is shown in figure 2.3, with gray lines illustrating the bins. The percentage of events in each $\not{E}_{\mathrm{T}}$ bin after the baseline selection is shown in table 2.2.

Leptonic decays of W bosons in which the lepton is not reconstructed constitute the majority of the $\mathrm{t\bar{t}}$ background. In these cases, $\not{E}_{\mathrm{T}}$ is primarily from the neutrino in the W decay (and possibly the missed lepton). Therefore the transverse mass of $\not{E}_{\mathrm{T}}$ and the b-jet from a given top decay has an endpoint at the top-quark mass. Thus, the $\mathrm{t\bar{t}}$ background can be significantly reduced by

	$t\bar{t}$/W+jets		$Z \to \nu\nu$		QCD		ttZ		T2tt(700,100)		T2tb(700,100)		T2tt(250,150)	
$\not{E}_T$ 250-300	339	58%	53	38%	15.2	48%	5.3	39%	4.0	12%	3.5	12%	23	40%
$\not{E}_T$ 300-400	191	33%	51	37%	12.2	39%	5.2	38%	8.8	26%	7.6	27%	22	39%
$\not{E}_T$ 400-500	37	6%	19.1	14%	1.7	5%	1.9	14%	8.7	26%	7.2	26%	7.0	13%
$\not{E}_T$ 500-600	9.0	2%	7.8	6%	1.8	6%	0.68	5%	6.5	20%	5.3	19%	2.7	5%
$\not{E}_T \geq$ 600	4.3	1%	7.1	5%	0.62	2%	0.51	4%	5.3	16%	4.4	16%	1.8	3%

Table 2.2: Event distribution across the $\not{E}_T$ bins after the baseline selection. The percent of events in each bin is relative to the baseline selection.

requiring $M_T(b_{1,2}, \not{E}_T) \geq 175$ for

$$M_T(b_{1,2}, \not{E}_T) = \min[M_T(b_1, \not{E}_T), M_T(b_2, \not{E}_T)]$$

where b_1 and b_2 are the jets with the highest and second highest CSV discriminator values respectively and

$$M_T(b_i, \not{E}_T) = \sqrt{2 \cdot p_T(b_i) \cdot \not{E}_T \cdot (1 - \cos\Delta\phi)} \ .$$

The minimum M_T between the two leading CSV jets is used since either of the two b-jets could be from a leptonic W boson decay. However, for some signal points a significant portion of the signal would be lost by this requirement. Therefore, the search region is split into two bins with $M_T(b_{1,2}, \not{E}_T) < 175$ and $M_T(b_{1,2}, \not{E}_T) \geq 175$. The separation power of $M_T(b_{1,2}, \not{E}_T)$ for some signal points and the need to keep the low $M_T(b_{1,2}, \not{E}_T)$ region for others is shown in the left plot of figure 2.7.

In the high $M_T(b_{1,2}, \not{E}_T)$ region requiring a reconstructed top-quark decay, as described in section 2.2.5, results in a very high signal purity for some models. However, as can be seen in the right plot of figure 2.7, even signal models with a relatively high efficiency for this region will not have a reconstructed top-quark for the majority of events. Thus the high $M_T(b_{1,2}, \not{E}_T)$ region is subdivided into

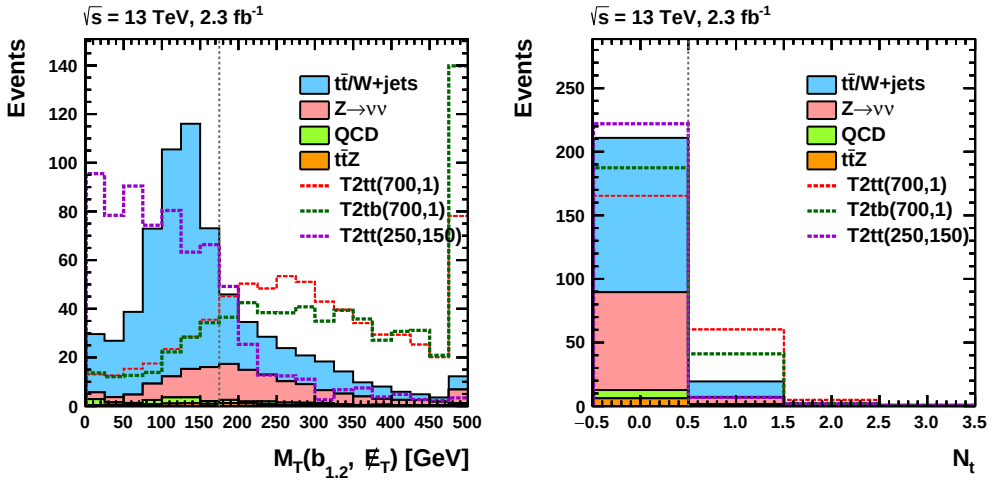

Figure 2.7: The $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$ (left) and N_{t} (right) distributions after the full baseline selection for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The N_{t} distribution also requires $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) > 175$ GeV. The signals are each normalized to the total background and the gray lines show the binning.

regions with zero and at least one reconstructed top decay.

The choice to allow events with only five jets into the search region is supplemented by binning in the number of jets. This helps to further discriminate against background events with fewer jets without loosing signal events with unreconstructed jets. Even though stop pair decays only produce six jets themselves, additional jets can be produced through initial state radiation. This can make the total number of jets in an event quite high, as can be seen in the plot on the right side of figure 2.8. Therefore bins are defined as a medium N_{J} region with five or six jets and a high N_{J} region with at least seven jets. Due to the already limited statistics in the high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$ region with one top, it is not further binned in N_{J}.

Events are binned in the number of medium b-jets to supplement the baseline selection on b-jets: $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$. This is shown in the left plot of figure 2.8 where there is a considerable difference between the signal and background distri-

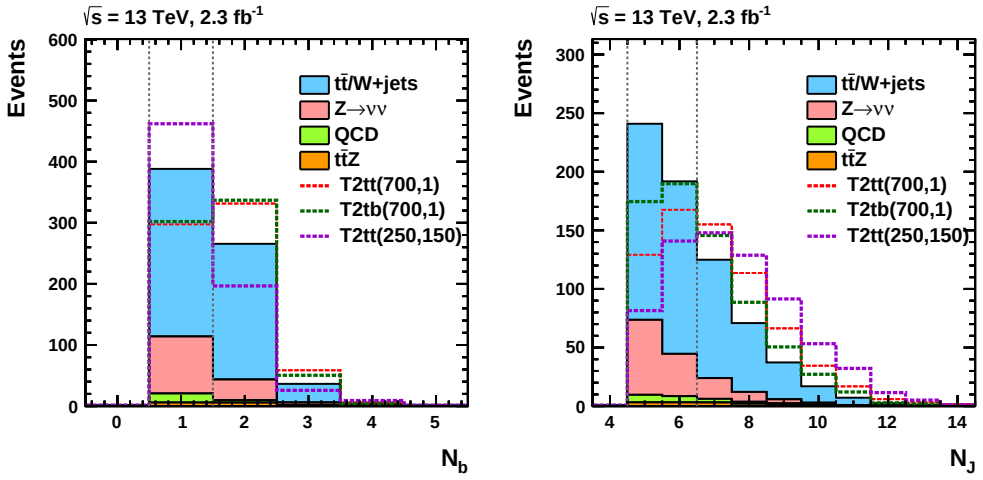

Figure 2.8: Distributions of N_{b} (left) and N_{J} (right) after the full baseline selection for the backgrounds (solid stacked histograms) and three example signal points (dotted color lines). The signals are each normalized to the total background and the gray lines show the binning.

	$t\bar{t}$/W+jets		$Z \to \nu\nu$		QCD		ttZ		T2tt(700,100)		T2tb(700,100)		T2tt(250,150)	
$N_{\mathrm{b}} = 1$														
low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}	113	19%	28	20%	6.0	19%	1.1	8%	1.1	3%	1.1	4%	8.9	16%
low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}	81	14%	10.5	8%	6.0	19%	1.4	10%	1.9	6%	1.4	5%	20.0	36%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}	55	9%	41	30%	7.8	25%	1.8	13%	3.5	10%	4.3	15%	3.0	5%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}	28	5%	11.9	9%	3.9	12%	1.2	9%	4.0	12%	3.1	11%	4.5	8%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{t}	7.3	1%	3.3	2%	0.77	2%	0.42	3%	2.9	9%	1.7	6%	0.36	1%
$N_{\mathrm{b}} \geq 2$														
low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}	248	43%	40	29%	8.6	27%	2.5	18%	2.5	7%	2.4	9%	13.0	23%
low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}	187	32%	15.0	11%	8.1	26%	3.4	25%	4.6	14%	3.1	11%	31	55%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}	86	15%	60	44%	9.0	28%	3.8	28%	8.4	25%	10.1	36%	4.7	8%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}	47	8%	17.6	13%	4.8	15%	3.0	22%	10.4	31%	8.0	29%	6.7	12%
high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{t}	12.7	2%	4.8	3%	1.1	3%	0.96	7%	7.5	22%	4.3	15%	0.56	1%

Table 2.3: Event distribution across the bins in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, N_{t}, and N_{b}. The percent of events in each bin is relative to the baseline selection.

butions for some signal points.

Table 2.3 shows the distribution of events across the bins in N_{b}, $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, and N_{t}. Bringing everything together, the baseline search region is subdivided into 50 disjoint bins. These bins cover the entire region and are summarized below:

- 5 divisions in $\not{E}_{\mathrm{T}}$: $[250, 300)$, $[300, 400)$, $[400, 500)$, $[500, 600)$, ≥ 600 GeV

- 2 divisions in N_{b}: 1, ≥ 2
- 5 interdependent divisions in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, and N_{t}:
  - low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \in [5, 6]$
  - low $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \geq 7$
  - high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, low N_{J}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \in [5, 6]$, $N_{\mathrm{t}} = 0$
  - high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{J}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \geq 7$, $N_{\mathrm{t}} = 0$
  - high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{t}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} \geq 1$.

Chapter 3

Background Estimation

All but one of the significant backgrounds are estimated by using MC to extrapolate their contributions to every bin of the search region from dedicated data control regions while the last is estimated directly from MC. A control region is defined for each background that is kinematically similar to but does not overlap with the search region. Each region is also designed to have a high purity for the background process it is being used to estimate and as many events as possible. These requirements result in regions that can propagate rare effects to the search region prediction as well as limit the statistical uncertainty of the prediction.

A transfer factor is then defined as the ratio of the number of MC events in the search region over the number of MC events in the control region. This is used to extrapolate from the number of data events observed in the control region to a prediction for the expected number of events in the search region. This method has the advantage of at least partially canceling many uncertainties that are shared between the control and search regions for a given process. In order to be confident in the use of MC to extrapolate between the search and control regions various corrections are applied. This ensures that the data is well-modeled

by the MC for all important features, as discussed in section 3.5.

In order to gain confidence in the background prediction methods the full prediction is run for additional control regions. These regions are distinct from both the search region and all control regions used to predict the various backgrounds. The outcome of these tests are shown in section 3.6.

3.1 Lost Leptons

A failure to identify the lepton from a leptonic W decay is the dominant source of SM background events in the majority of the search region bins. This happens either because the lepton is outside of kinematic acceptance (i.e. $p_{\mathrm{T}} < 5$ GeV or $|\eta| > 2.4$) or does not meet the identification and isolation requirements imposed by the lepton vetoes. The associated neutrino can generate the large $\not{E}_{\mathrm{T}}$ necessary for these events to enter the search region. This background primarily consists of $t\bar{t}$ and W+jets events with ttW and single-top events also contributing a small amount. The majority of this background (up to 80%) comes from events where leptons are reconstructed as jets rather than identified by the lepton vetoes.

In order to estimate this Lost Lepton Background (LLB) background a single-lepton control region is defined by inverting the electron and muon vetoes to require at least one isolated "veto" electron or muon. The remaining baseline selection is left as is. The transverse mass between the lepton and $\not{E}_{\mathrm{T}}$ ($M_{\mathrm{T}}(\ell, \not{E}_{\mathrm{T}})$) is required to be below 100 GeV in order to suppress potential signal contamination. This also ensures that this control region is orthogonal to the search region used in the single-lepton stop search in order to facilitate an eventual combination of the results. If multiple leptons in an event meet the veto selection requirements one is randomly chosen to calculate $M_{\mathrm{T}}(\ell, \not{E}_{\mathrm{T}})$.

The binning used for the search region is also applied to this control region with the following exceptions. The $\not{E}_{\mathrm{T}}$ distributions from events with lost leptons only have a small dependence on the number of b-tagged jets within the current statistical uncertainties of the MC. Therefore the bins in N_{b} can be combined to simply require $N_{\mathrm{b}} \geq 1$. This results in a final control region where each bin has two to four times more events than the corresponding bins in the search region. All five bins in $\not{E}_{\mathrm{T}}$ are also combined in the high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, high N_{t} region to compensate for the very low statistics. This is feasible since the transfer factors (as defined in the next paragraph) in the high $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, zero top bins show no dependence on $\not{E}_{\mathrm{T}}$. A systematic uncertainty is placed on this extrapolation based on the relative uncertainty in each $\not{E}_{\mathrm{T}}$ bin with $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) > 175$ and $N_{\mathrm{t}} = 0$. These mergers result in the single-lepton control region being divided into a total of 21 bins. Figure 3.1 shows a comparison between data and MC across all the single-lepton control region bins. The event counts in both data and MC are also given in table 3.1.

The number of data events observed in each bin in the control region is translated to the corresponding bins in the search region by using a transfer factor derived from MC. For each search region bin, the transfer factor is defined as:

$$TF_{\mathrm{LLB}} = \frac{N_{\mathrm{MC}}(0l)}{N_{\mathrm{MC}}(1l)}$$

where $N_{\mathrm{MC}}(0l)$ and $N_{\mathrm{MC}}(1l)$ are the number of events in MC for the search region bin and its corresponding control region bin respectively. The final prediction is then given by:

$$N_{\mathrm{pred}}^{\mathrm{LLB}} = TF_{\mathrm{LLB}} \cdot N_{\mathrm{data}}(1l)$$

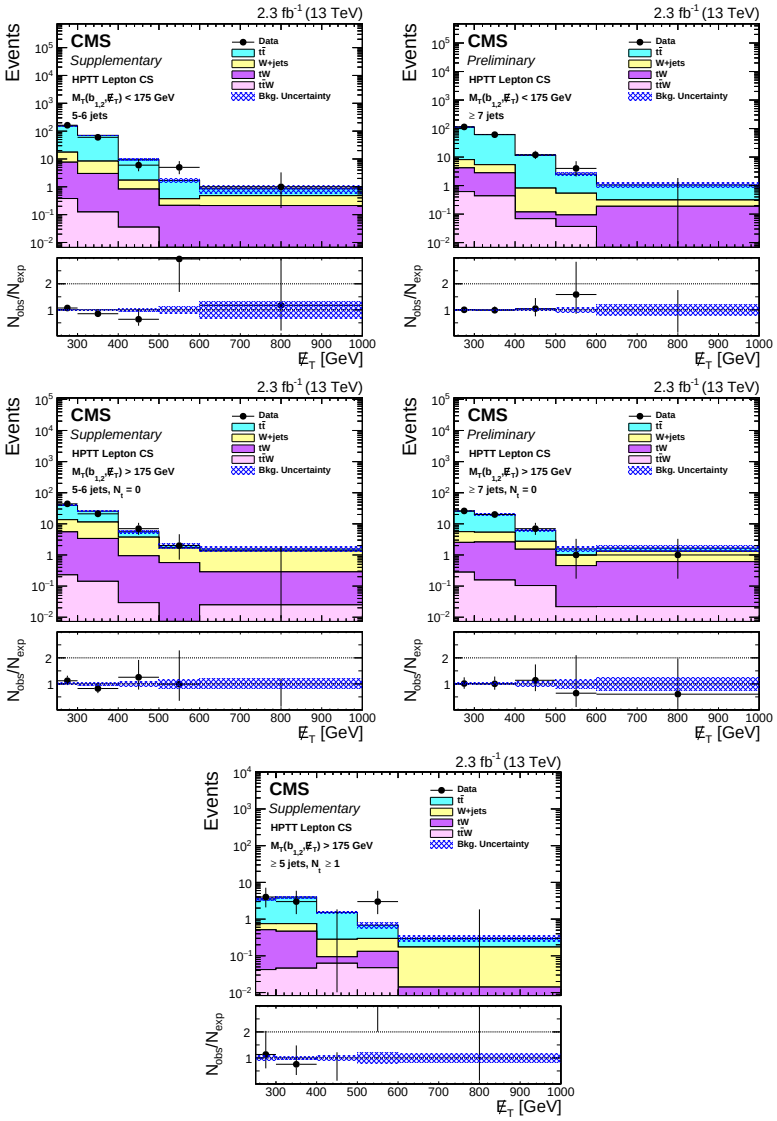

Figure 3.1: Data-MC comparisons across all bins in the single-lepton control region. The black points and stacked histograms are data and the contributing samples in MC respectively. In the data-over-MC ratio plot below each histogram, the error bars and blue hatched band are the statistical uncertainties on data and MC respectively. Each plot shows the $\not{E}_{\mathrm{T}}$ bins for the given bin in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, and N_{t}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \in [5,6]$ (top left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \geq 7$ (top right); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \in [5,6]$, $N_{\mathrm{t}} = 0$ (middle left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \geq 7$, $N_{\mathrm{t}} = 0$ (middle right); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} \geq 1$ (bottom). Although the $\not{E}_{\mathrm{T}}$ bins are shown in the bottom plot they are integrated over for the LLB prediction.

$\not{E}_{\mathrm{T}}$[GeV]	N_{data}	$\mathrm{t\bar{t}}$	W+jets	tW	ttW	ttZ	Total Bkg
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$ GeV, $5-6$ jets							
250-300	159	168.68 ± 1.96	12.32 ± 1.42	9.07 ± 1.32	0.47 ± 0.07	0.20 ± 0.10	190.28 ± 2.75
300-400	57	77.82 ± 1.24	6.68 ± 0.98	3.61 ± 0.84	0.16 ± 0.05	0.22 ± 0.10	88.34 ± 1.79
400-500	6	9.64 ± 0.44	1.13 ± 0.33	0.99 ± 0.45	0.04 ± 0.02	0.00 ± 0.00	11.76 ± 0.71
500-600	2	1.65 ± 0.18	0.20 ± 0.07	0.26 ± 0.22	0.00 ± 0.00	0.00 ± 0.00	2.11 ± 0.29
>600	1	0.47 ± 0.09	0.33 ± 0.24	0.27 ± 0.23	-0.00 ± 0.00	0.00 ± 0.00	1.07 ± 0.35
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$ GeV, ≥ 7 jets							
250-300	108	152.50 ± 1.75	5.69 ± 0.80	5.13 ± 1.01	0.89 ± 0.10	0.19 ± 0.10	163.51 ± 2.17
300-400	59	81.49 ± 1.27	3.75 ± 0.54	3.41 ± 0.78	0.63 ± 0.09	0.01 ± 0.01	88.65 ± 1.58
400-500	12	15.37 ± 0.55	1.00 ± 0.19	0.07 ± 0.07	0.10 ± 0.04	0.04 ± 0.04	16.48 ± 0.59
500-600	3	2.85 ± 0.23	0.64 ± 0.25	0.08 ± 0.08	0.05 ± 0.03	0.00 ± 0.00	3.58 ± 0.35
>600	0	1.06 ± 0.15	0.18 ± 0.06	0.27 ± 0.27	0.00 ± 0.01	0.00 ± 0.00	1.51 ± 0.31
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) > 175$ GeV, $N_{\mathrm{t}} = 0$, $5-6$ jets							
250-300	36	38.32 ± 0.89	12.04 ± 1.45	7.88 ± 1.22	0.34 ± 0.06	0.35 ± 0.16	58.59 ± 2.10
300-400	21	21.10 ± 0.66	11.94 ± 1.39	4.84 ± 0.94	0.21 ± 0.05	0.20 ± 0.13	38.08 ± 1.81
400-500	7	2.77 ± 0.24	4.15 ± 0.58	1.36 ± 0.45	0.04 ± 0.03	0.00 ± 0.00	8.27 ± 0.77
500-600	2	0.52 ± 0.11	1.65 ± 0.30	0.84 ± 0.42	0.01 ± 0.02	0.00 ± 0.00	3.01 ± 0.53
>600	0	0.30 ± 0.09	1.58 ± 0.34	0.39 ± 0.28	0.04 ± 0.02	0.01 ± 0.01	2.28 ± 0.45
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) > 175$ GeV, $N_{\mathrm{t}} = 0$, ≥ 7 jets							
250-300	26	27.43 ± 0.75	4.20 ± 0.60	3.08 ± 0.74	0.38 ± 0.07	0.04 ± 0.03	34.74 ± 1.21
300-400	19	19.97 ± 0.64	3.78 ± 0.41	3.34 ± 0.80	0.21 ± 0.07	0.07 ± 0.05	27.17 ± 1.11
400-500	6	4.67 ± 0.31	1.66 ± 0.22	1.94 ± 0.60	0.14 ± 0.05	0.03 ± 0.03	8.30 ± 0.71
500-600	2	0.79 ± 0.14	0.72 ± 0.16	0.59 ± 0.33	0.03 ± 0.02	0.00 ± 0.00	2.10 ± 0.40
>600	0	0.46 ± 0.10	0.96 ± 0.31	0.81 ± 0.43	0.03 ± 0.03	0.01 ± 0.01	2.24 ± 0.54
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) > 175$ GeV, $N_{\mathrm{t}} \geq 1$, ≥ 5 jets							
>250	9	14.42 ± 0.57	1.93 ± 0.22	1.89 ± 0.59	0.38 ± 0.07	0.00 ± 0.00	18.25 ± 0.85

Table 3.1: Observed events in data and MC in the single-lepton control region. The uncertainties are statistical only.

where $N_{\mathrm{data}}(1l)$ is the number of events observed in data in the appropriate control region bin. The transfer factor and prediction for each bin is given in table 3.2.

Although the $M_{\mathrm{T}}(\ell, \not{E}_{\mathrm{T}}) < 100$ GeV cut reduces the potential signal contamination to a negligible level for signal points with a high stop mass it can still be significant for other points. Any remaining signal contamination in this control region is taken into account in the final result by simultaneously fitting the single lepton control regions with the search regions. See section 4.1 for more information.

$\not\!\!E_{\mathrm{T}}$[GeV]	N_{data}	TF_{LLB}	$N_{\mathrm{pred}}^{\mathrm{LLB}}$	TF_{LLB}	$N_{\mathrm{pred}}^{\mathrm{LLB}}$
		$N_{\mathrm{b}} = 1$		$N_{\mathrm{b}} \geq 2$	
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) < 175$ GeV, $5-6$ jets					
250−300	159	0.37 ± 0.01	58.30 ± 5.00 (stat.) ± 2.92 (syst.)	0.38 ± 0.01	59.72 ± 4.97 (stat.) ± 2.46 (syst.)
300−400	57	0.38 ± 0.02	21.49 ± 3.00 (stat.) ± 1.16 (syst.)	0.40 ± 0.01	23.05 ± 3.15 (stat.) ± 0.92 (syst.)
400−500	6	0.41 ± 0.04	2.45 ± 1.03 (stat.) ± 0.17 (syst.)	0.46 ± 0.05	2.78 ± 1.17 (stat.) ± 0.19 (syst.)
500−600	2	0.38 ± 0.08	0.76 ± 0.56 (stat.) ± 0.07 (syst.)	0.35 ± 0.08	0.70 ± 0.52 (stat.) ± 0.05 (syst.)
>600	1	0.28 ± 0.12	0.28 ± 0.30 (stat.) ± 0.04 (syst.)	0.38 ± 0.15	0.38 ± 0.40 (stat.) ± 0.05 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) < 175$ GeV, ≥ 7 jets					
250−300	108	0.27 ± 0.01	28.71 ± 2.91 (stat.) ± 2.01 (syst.)	0.31 ± 0.01	33.73 ± 3.38 (stat.) ± 2.00 (syst.)
300−400	59	0.28 ± 0.01	16.78 ± 2.30 (stat.) ± 1.24 (syst.)	0.33 ± 0.01	19.52 ± 2.63 (stat.) ± 1.10 (syst.)
400−500	12	0.24 ± 0.02	2.87 ± 0.87 (stat.) ± 0.22 (syst.)	0.38 ± 0.03	4.54 ± 1.37 (stat.) ± 0.30 (syst.)
500−600	3	0.31 ± 0.06	0.94 ± 0.57 (stat.) ± 0.10 (syst.)	0.38 ± 0.06	1.13 ± 0.68 (stat.) ± 0.11 (syst.)
>600	0	0.31 ± 0.10	0.00 ± 0.56 (stat.) ± 0.00 (syst.)	0.33 ± 0.09	0.00 ± 0.59 (stat.) ± 0.00 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} = 0$, $5-6$ jets					
250−300	36	0.45 ± 0.03	16.29 ± 2.90 (stat.) ± 0.99 (syst.)	0.25 ± 0.02	8.99 ± 1.66 (stat.) ± 0.54 (syst.)
300−400	21	0.46 ± 0.03	9.59 ± 2.21 (stat.) ± 0.70 (syst.)	0.23 ± 0.02	4.88 ± 1.16 (stat.) ± 0.31 (syst.)
400−500	7	0.63 ± 0.11	4.44 ± 1.85 (stat.) ± 0.24 (syst.)	0.23 ± 0.04	1.62 ± 0.68 (stat.) ± 0.16 (syst.)
500−600	2	0.41 ± 0.12	0.82 ± 0.63 (stat.) ± 0.07 (syst.)	0.15 ± 0.05	0.29 ± 0.23 (stat.) ± 0.06 (syst.)
>600	0	0.22 ± 0.06	0.00 ± 0.40 (stat.) ± 0.00 (syst.)	0.27 ± 0.15	0.00 ± 0.49 (stat.) ± 0.00 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} = 0$, ≥ 7 jets					
250−300	26	0.34 ± 0.02	8.84 ± 1.83 (stat.) ± 0.54 (syst.)	0.23 ± 0.02	5.90 ± 1.27 (stat.) ± 0.46 (syst.)
300−400	19	0.36 ± 0.03	6.77 ± 1.67 (stat.) ± 0.46 (syst.)	0.19 ± 0.02	3.63 ± 0.89 (stat.) ± 0.27 (syst.)
400−500	6	0.29 ± 0.05	1.72 ± 0.76 (stat.) ± 0.11 (syst.)	0.21 ± 0.04	1.25 ± 0.58 (stat.) ± 0.10 (syst.)
500−600	2	0.38 ± 0.11	0.75 ± 0.58 (stat.) ± 0.11 (syst.)	0.22 ± 0.12	0.43 ± 0.38 (stat.) ± 0.08 (syst.)
>600	0	0.28 ± 0.16	0.00 ± 0.51 (stat.) ± 0.04 (syst.)	0.06 ± 0.03	0.00 ± 0.11 (stat.) ± 0.00 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} \geq 1$, ≥ 5 jets					
250−300	9	0.12 ± 0.02	1.10 ± 0.40 (stat.) ± 0.16 (syst.)	0.07 ± 0.01	0.60 ± 0.21 (stat.) ± 0.09 (syst.)
300−400	9	0.15 ± 0.03	1.38 ± 0.54 (stat.) ± 0.29 (syst.)	0.09 ± 0.01	0.83 ± 0.30 (stat.) ± 0.15 (syst.)
400−500	9	0.07 ± 0.01	0.66 ± 0.24 (stat.) ± 0.21 (syst.)	0.03 ± 0.01	0.28 ± 0.11 (stat.) ± 0.08 (syst.)
500−600	9	0.02 ± 0.01	0.22 ± 0.09 (stat.) ± 0.12 (syst.)	0.03 ± 0.02	0.27 ± 0.17 (stat.) ± 0.15 (syst.)
−600	9	0.02 ± 0.00	0.14 ± 0.06 (stat.) ± 0.07 (syst.)	0.01 ± 0.01	0.12 ± 0.07 (stat.) ± 0.06 (syst.)

Table 3.2: The transfer factor and prediction for the LLB in each search region bin. The number of events observed in data is given for the corresponding single lepton control region bins.

Uncertainties

The dominant uncertainty for the LLB prediction is from the limited statistics of the control region and ranges from 8% in the more populated bins to 100% in the tightest bins. The MC statistical uncertainties, which have values ranging from 2-57% and 1-34% for the control and search regions respectively, also have a considerable effect in some bins.

Corrections for differences in the efficiency of identifying leptons between data and MC are applied to ensure a good description of both the search and control regions. These are especially important for this prediction since the effect is anti-

correlated between the control region where leptons are selected and the search region where they are vetoed. These uncertainties range from 1-7%. A more detailed discussion is in section 3.5.2.

Possible differences in the background composition between data and MC are evaluated by independently varying the W+jets and $t\bar{t}$ cross sections by 20%. This results in uncertainties ranging from 0-11% which are assigned to each bin. An uncertainty of 10% is assigned to $t\bar{t}$ and non-$t\bar{t}$ processes in the one top bins to account for differences in the top-tagging efficiency and mis-tagging rate respectively. More information is in section 3.5.5. Standard systematic uncertainties for pileup reweighting (0-19%), jet energy scales (0-11%), b-tagging (0-7%), and the tau veto (1-11%) are also applied. They are discussed further in section 3.5.

3.2 Invisible Z Decays

The second largest source of background events in the search region is events in which a Z boson decays into neutrinos. The neutrinos manifest as missing energy that can allow these events to enter the search region. There are two potential choices for a control region for this process: $Z \to \ell\ell$+jets and γ+jets.

The benefit of using the $Z \to \ell\ell$+jets control region is that both the $Z \to \ell\ell$ and $Z \to \nu\nu$ processes have very similar kinematics. However, this region is very limited in statistics especially in the tight search region used in this analysis. On the other hand, the γ+jets control region has a larger cross section than $Z \to \ell\ell$+jets by at least a factor of five as well as similar leading order Feynman diagram to $Z \to \nu\nu$+jets. The drawback of this region is that different quark-boson couplings, the masslessness of the photon, and other effects related to fake and fragmentation photons in the γ+jets sample need to be accounted for in order

to use it for a $Z \to \nu\nu$ prediction. Rather than just choosing one region, a hybrid method is used which takes advantage of the strengths of both regions. First, the $Z \to \ell\ell$+jets control region is used to correct the overall normalization of the $Z \to \nu\nu$ MC. Then the γ+jets control region is used for more finely binned shape corrections.

In order to emulate the $Z \to \nu\nu$ process the photon or dilepton system is removed from an event before calculating $\not{E}_T$ and any variables dependent on it. $\not{E}_T^{\gamma}$ and $\not{E}_T^{\ell\ell}$ represent this modified $\not{E}_T$ after the photon and dilepton system is removed respectively. The final prediction for the number of $Z \to \nu\nu$ events in each bin of the search region is given by:

$$N_{\text{pred}}^{Z\to\nu\nu} = N_{\text{MC}}^{Z\to\nu\nu} \cdot R_Z \cdot S_\gamma$$

where $N_{\text{MC}}^{Z\to\nu\nu}$ is the number of $Z \to \nu\nu$ events in MC, R_Z is the normalization correction, and S_γ is the shape correction.

Normalization Correction

The normalization correction R_Z is defined as the ratio of data over MC events in the $Z \to \ell\ell$ control region after accounting for any contamination from other processes. The definition of this control region starts with a loose version of baseline selection. Namely, removing the $\Delta\phi$ cuts between $\not{E}_T^{\ell\ell}$ and the four leading jets and lowering the $\not{E}_T^{\ell\ell}$ cut to 100 GeV. In order to select a high purity sample two same flavor opposite sign leptons are required. These leptons must pass a medium identification working point and have an invariant mass within about 10 GeV of the Z boson ($80 < M_{\ell\ell} < 100$ GeV).

Even within this tight mass window the contamination from $t\bar{t}$ events is not

negligible. A similar factor (R_{T}) is defined for $\mathrm{t\bar{t}}$ events to account for this. R_{T} is defined as the ratio of data over MC events outside of the Z-mass window ($20 < M_{\ell\ell} < 80$ or > 100 GeV) after accounting for contamination from other processes. The following set of equations is solved to take the cross-contamination into account when obtaining R_{Z} and R_{T}:

$$\begin{bmatrix} N_{\mathrm{data}}(\text{on Z}) \\ N_{\mathrm{data}}(\text{off Z}) \end{bmatrix} = \begin{pmatrix} N_{\mathrm{MC}}^{\mathrm{Z}\to\ell\ell}(\text{on Z}) & N_{\mathrm{MC}}^{\mathrm{t\bar{t}}}(\text{on Z}) \\ N_{\mathrm{MC}}^{\mathrm{Z}\to\ell\ell}(\text{off Z}) & N_{\mathrm{MC}}^{\mathrm{t\bar{t}}}(\text{off Z}) \end{pmatrix} \cdot \begin{bmatrix} R_{\mathrm{Z}} \\ R_{\mathrm{T}} \end{bmatrix}.$$

The small contributions from tZ and ttZ are included with $\mathrm{Z} \to \ell\ell$ and the small contributions from tW and ttW are included with $\mathrm{t\bar{t}}$. R_{Z} and R_{T} are calculated separately for $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$ in order to account for any effects from heavy flavor production.

To check the dependence of R_{Z} on $\not\!\!E_{\mathrm{T}}^{\ell\ell}$ it is calculated both for all events with $\not\!\!E_{\mathrm{T}}^{\ell\ell} > 100$ GeV and also in bins of $\not\!\!E_{\mathrm{T}}^{\ell\ell}$. The constant value and the linear fit of the binned values are seen to be constant within statistical uncertainties. Therefore lowering the $\not\!\!E_{\mathrm{T}}^{\ell\ell}$ cut to 100 GeV allows the statistical uncertainty of this correction to be decreased without introducing any statistically significant biases. The difference between the constant value and the linear fit is taken as a systematic uncertainty on the $\not\!\!E_{\mathrm{T}}^{\ell\ell}$ extrapolation. This is also done separately for the two N_{b} cases.

The measured values of R_{Z} for the $N_{\mathrm{b}} = 1$ and $N_{\mathrm{b}} \geq 2$ regions are $0.94 \pm 0.12(\text{stat.}) \pm 0.06(\not\!\!E_{\mathrm{T}}^{\ell\ell})$ and $0.84 \pm 0.19(\text{stat.}) \pm 0.01(\not\!\!E_{\mathrm{T}}^{\ell\ell})$ respectively, where the second ($\not\!\!E_{\mathrm{T}}^{\ell\ell}$) errors are from the extrapolation in $\not\!\!E_{\mathrm{T}}^{\ell\ell}$.

Shape Correction

There are three different types of photons in the γ+jets control region: prompt photons that are produced directly, prompt photons that are produced through fragmentation, and fakes. Since the latter two have no analog for $Z \to \nu\nu$ the impact of any mis-modeling needs to be understood. Prompt photons are reconstructed photons that correspond to a real photon. For the sake of this search photons are considered to be prompt when they are within $\Delta R(\gamma_{\text{gen}}, \gamma_{\text{reco}}) < 0.1$ and $0.5 < p_{\text{T}}(\gamma_{\text{gen}})/p_{\text{T}}(\gamma_{\text{reco}}) < 2$ of a generator-level photon. Fakes are any reconstructed photons that are not matched to a generator-level photon. Fragmentation photons are produced during the hadronization of quarks. To differentiate these from direct photons every prompt photon is checked to see if it is within $\Delta R(\gamma, \text{parton}) = 0.4$ of a generator level quark or gluon. If so it is considered to be a fragmentation photon. Otherwise it is considered to be a direct photon. The γ+jets and QCD MC samples are combined in order to study all three of these components. In order to avoid double counting events during this combination only direct photons are used from the γ+jets MC and only fragmentation and fake photons are used from the QCD MC.

Since there are no similar processes for $Z \to \nu\nu$ any mis-modelings of the fragmentation and fake components can impact the final background estimation if the shapes of observables are severely affected. For fragmentation photons it is not possible to distinguish them from direct photons in data. Therefore only the impact of mis-modeling the fragmentation component can be checked. This is done by scaling its contribution in MC by $\pm 50\%$ and evaluating the effect on the $\not{E}_{\text{T}}^{\gamma}$ shapes. This scaling has less than a 5% effect on the $\not{E}_{\text{T}}^{\gamma}$ shapes. The fraction of fake photons in data is estimated using a maximum likelihood fit and seen to

be consistent with MC. Therefore the same ±50% scaling procedure is repeated for fake photons. This effect on the $\not\!\!E_{\mathrm{T}}^{\gamma}$ shapes is also seen to be small. Since the γ+jets control region is only used for shape corrections these tests are sufficient to gain confidence that any mis-modeling effects will not significantly affect the final prediction.

ttγ events can also contribute to the γ+jets sample. This component is taken directly from MC after repeating the above scaling procedure to ensure that there is no significant effect on the $\not\!\!E_{\mathrm{T}}^{\gamma}$ shapes.

The γ+jets control region is defined by first requiring $\not\!\!E_{\mathrm{T}} < 200$ GeV in order to suppress potential signal contamination. Then $\not\!\!E_{\mathrm{T}}^{\gamma}$ is used in place of $\not\!\!E_{\mathrm{T}}$ for both the baseline selection and the binning. Since this is a shape correction the γ+jets MC is normalized to the data across the full control region. As with the LLB prediction, the bins in N_{b} are combined in order to increase the available statistics after ensuring this does not introduce any bias in the $\not\!\!E_{\mathrm{T}}^{\gamma}$ shapes within statistical uncertainties. Then S_{γ} is calculated separately for each bin as:

$$S_{\gamma} = \frac{N_{\mathrm{data}}(\gamma + \mathrm{jets})}{N_{\mathrm{MC}}(\gamma + \mathrm{jets})}.$$

Figure 3.2 shows a comparison of data and MC across all the single-photon control region bins. The data-over-MC ratios in each bin are the S_{γ} factors used for the $\mathrm{Z} \to \nu\nu$ prediction for the corresponding search region bins. The $\mathrm{Z} \to \nu\nu$ prediction for each bin of the search region is given in table 3.3. The S_{γ} factors and raw MC yield are also included.

To verify the assumption that the shape differences between data and MC are similar for both $\mathrm{Z} \to \nu\nu$ and γ+jets, the ratios of data over MC for $\mathrm{Z} \to \ell\ell$+jets and γ+jets are compared to each other. This is done in two regions. In

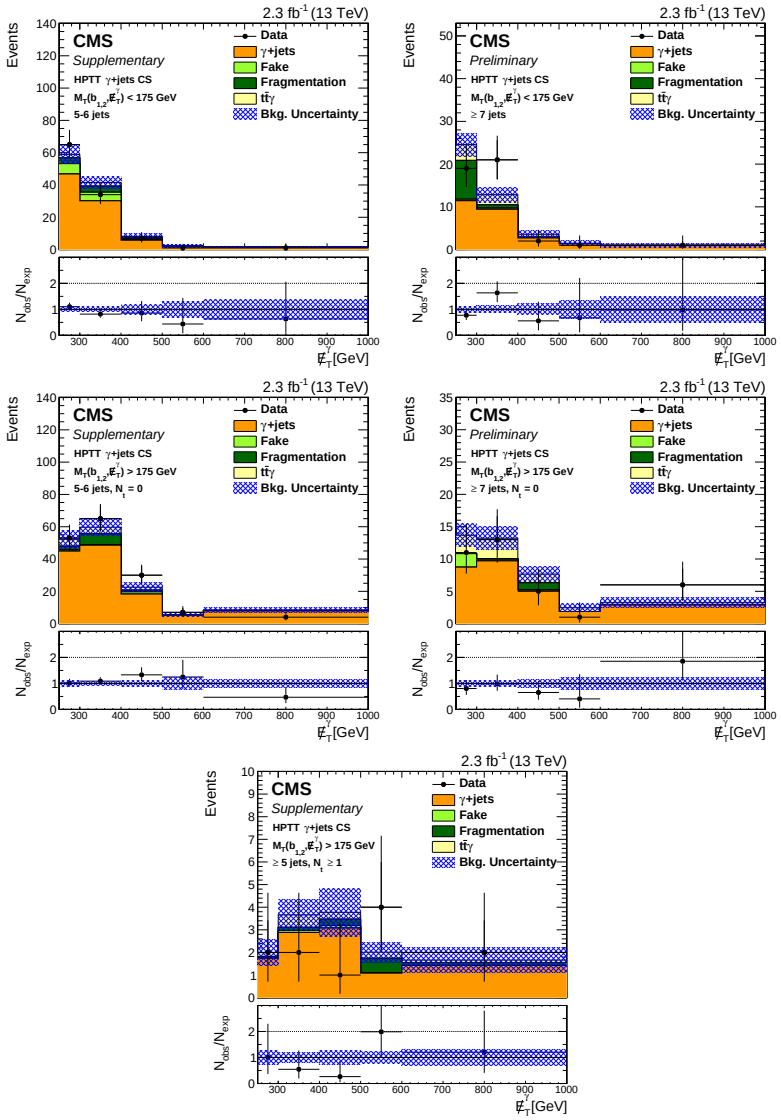

Figure 3.2: Data-MC comparisons across all bins in the γ+jets control region. The black points and stacked histograms are data and the contributing samples in MC respectively. In the data-over-MC ratio plot below each histogram, the black error bars and blue hatched band are the statistical uncertainties on data and MC respectively. Each plot shows the $\not{E}_{\mathrm{T}}$ bins for the given bin in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, and N_{t}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \in [5, 6]$ (top left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \geq 7$ (top right); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \in [5, 6]$, $N_{\mathrm{t}} = 0$ (middle left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \geq 7$, $N_{\mathrm{t}} = 0$ (middle right); and $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} \geq 1$ (bottom). The data-over-MC ratios in each bin are the S_{γ} factors used for the $\mathrm{Z} \to \nu\nu$ prediction.

$\not\!\!E_{\mathrm{T}}$[GeV]	S_γ	$N_{\mathrm{MC}}^{\mathrm{Z}\to\nu\nu}$	$N_{\mathrm{pred}}^{\mathrm{Z}\to\nu\nu}$	$N_{\mathrm{MC}}^{\mathrm{Z}\to\nu\nu}$	$N_{\mathrm{pred}}^{\mathrm{Z}\to\nu\nu}$
		$N_{\mathrm{b}}=1$		$N_{\mathrm{b}}\geq 2$	
$M_{\mathrm{T}}(\mathrm{b}_{1,2},\not\!\!E_{\mathrm{T}})<175$ GeV, $5-6$ jets, $N_{\mathrm{t}}\geq 0$					
250−300	1.02±0.17	13.22±0.98	13.12±2.36 (stat.) ±2.32 (syst.)	4.95±0.52	4.51± 0.88 (stat.) ±1.05 (syst.)
300−400	0.88±0.17	9.39±0.74	8.06±1.66 (stat.) ±1.39 (syst.)	4.22±0.48	3.32±0.74 (stat.) ±0.78 (syst.)
400−500	0.75±0.34	1.92±0.27	1.40±0.66 (stat.) ±0.32 (syst.)	0.83±0.17	0.56±0.28 (stat.) ±0.14 (syst.)
500−600	0.44±0.46	0.58±0.13	0.25±0.27 (stat.) ±0.07 (syst.)	0.35±0.13	0.14±0.15 (stat.) ±0.05 (syst.)
>600	0.64±0.68	0.21±0.06	0.13±0.15 (stat.) ±0.05 (syst.)	0.08±0.04	0.05±0.05 (stat.) ±0.02 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2},\not\!\!E_{\mathrm{T}})<175$ GeV, ≥ 7 jets, $N_{\mathrm{t}}\geq 0$					
250−300	0.78±0.20	3.98±0.49	3.05±0.86 (stat.) ±0.63 (syst.)	1.44±0.28	1.01±0.32 (stat.) ±0.24 (syst.)
300−400	1.66±0.43	2.92±0.38	4.73±1.37 (stat.) ±1.10 (syst.)	1.52±0.27	2.25±0.70 (stat.) ±0.57 (syst.)
400−500	0.57±0.42	1.52±0.25	0.84±0.64 (stat.) ±0.16 (syst.)	0.32±0.09	0.16±0.13 (stat.) ±0.04 (syst.)
500−600	0.68±0.73	0.14±0.06	0.09±0.11 (stat.) ±0.02 (syst.)	0.16±0.07	0.10±0.12 (stat.) ±0.03 (syst.)
>600	0.99±1.10	0.42±0.13	0.40±0.46 (stat.) ±0.08 (syst.)	0.11±0.05	0.10±0.12 (stat.) ±0.04 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2},\not\!\!E_{\mathrm{T}})\geq 175$ GeV, $5-6$ jets, $N_{\mathrm{t}}=0$					
250−300	1.05±0.18	12.48±0.93	12.79±2.42 (stat.) ±2.24 (syst.)	5.18±0.57	4.87±1.00 (stat.) ±1.15 (syst.)
300−400	1.10±0.17	16.38±1.04	17.68±2.91 (stat.) ±3.05 (syst.)	6.71±0.63	6.64±1.19 (stat.) ±1.53 (syst.)
400−500	1.21±0.28	6.72±0.58	7.97±1.96 (stat.) ±1.61 (syst.)	2.74±0.37	2.98±0.80 (stat.) ±0.74 (syst.)
500−600	1.27±0.55	3.14±0.36	3.89±1.75 (stat.) ±0.69 (syst.)	1.33±0.22	1.50±0.70 (stat.) ±0.37 (syst.)
>600	0.48±0.25	2.66±0.30	1.23±0.67 (stat.) ±0.22 (syst.)	0.79±0.15	0.34±0.19 (stat.) ±0.08 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2},\not\!\!E_{\mathrm{T}})\geq 175$ GeV, ≥ 7 jets, $N_{\mathrm{t}}=0$					
250−300	0.82±0.27	3.28±0.38	2.62±0.91 (stat.) ±0.50 (syst.)	1.72±0.27	1.26±0.46 (stat.) ±0.30 (syst.)
300−400	1.15±0.33	4.11±0.48	4.62±1.44 (stat.) ±0.97 (syst.)	1.90±0.31	1.96±0.65 (stat.) ±0.54 (syst.)
400−500	0.66±0.31	2.09±0.32	1.35±0.67 (stat.) ±0.25 (syst.)	0.85±0.16	0.51±0.26 (stat.) ±0.13 (syst.)
500−600	0.41±0.42	1.03±0.17	0.42±0.43 (stat.) ±0.09 (syst.)	0.31±0.11	0.12±0.13 (stat.) ±0.03 (syst.)
>600	1.88±0.90	1.24±0.20	2.27±1.15 (stat.) ±0.43 (syst.)	0.46±0.12	0.77±0.42 (stat.) ±0.21 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2},\not\!\!E_{\mathrm{T}})\geq 175$ GeV, ≥ 5 jets, $N_{\mathrm{t}}\geq 1$					
250−300	1.02±0.78	0.31±0.08	0.31±0.25 (stat.) ±0.07 (syst.)	0.12±0.04	0.11±0.09 (stat.) ±0.03 (syst.)
300−400	0.55±0.40	0.64±0.13	0.35±0.26 (stat.) ±0.07 (syst.)	0.25±0.07	0.12±0.10 (stat.) ±0.03 (syst.)
400−500	0.27±0.28	0.80±0.17	0.21±0.22 (stat.) ±0.4 (syst.)	0.15±0.05	0.04±0.04 (stat.) ±0.01 (syst.)
500−600	2.02±1.10	0.28±0.08	0.55±0.34 (stat.) ±0.10 (syst.)	0.18±0.06	0.32±0.20 (stat.) ±0.09 (syst.)
>600	1.22±0.95	0.51±0.11	0.61±0.49 (stat.) ±0.11 (syst.)	0.36±0.09	0.39±0.32 (stat.) ±0.10 (syst.)

Table 3.3: The $\mathrm{Z}\to\nu\nu$ prediction for each search region bin. The S_γ factors and raw MC yield are also included.

the first region, the N_{b} cut inverted to $N_{\mathrm{b}}=0$ and the $\Delta\phi$ cuts between the four leading jets and $\not\!\!E_{\mathrm{T}}^{\ell\ell}/\not\!\!E_{\mathrm{T}}^{\gamma}$ are removed so that a fairly fine binning in $\not\!\!E_{\mathrm{T}}^{\ell\ell}/\not\!\!E_{\mathrm{T}}^{\gamma}$can be used. In the second region, the full baseline selection is retained but a coarser binning in $\not\!\!E_{\mathrm{T}}^{\ell\ell}/\not\!\!E_{\mathrm{T}}^{\gamma}$ is used. Both processes agree within the available statistics as can be seen in figure 3.3.

Uncertainties

The dominant uncertainty for the $\mathrm{Z}\to\nu\nu$ prediction is from the statistical uncertainties on the data in the γ+jets control region with bins ranging from 12-100%. The MC statistical uncertainty also has a significant impact in the control

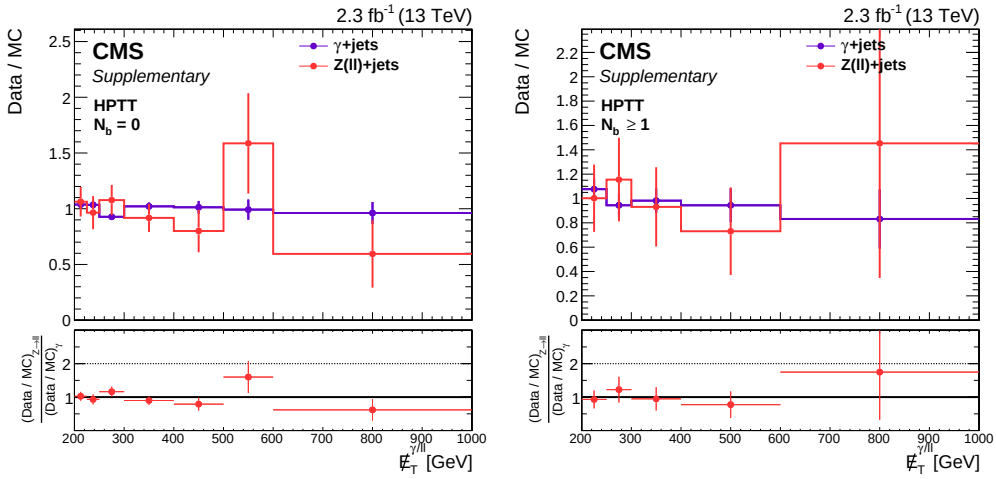

Figure 3.3: Comparison of the data over MC ratios for the $Z \to \ell\ell$ (red) and γ+jets (purple) processes in the relaxed region with the $N_b = 0$ and the $\Delta\phi$ cuts removed (left) and in the baseline selection (right). Both processes agree within statistical uncertainties.

and search regions with ranges 9-49% and 6-51% respectively.

The uncertainty on R_Z is 14% and 23% for the $N_b = 1$ and $N_b \geq 2$ regions respectively, and is actually the dominant uncertainty in some of the more populated bins. Standard corrections and systematic uncertainties for pileup reweighting (0-15%), jet energy scales (0-34%), b-tagging (0-9%), and the lepton vetoes (1-2%) are also applied. They are discussed further in section 3.5.

3.3 QCD Multijets

After the baseline selection requiring $\Delta\phi$ between $\not{E}_T$ and each of the four leading jets to greater than 0.5 the contribution from QCD multijet events is almost negligible in most of the search region bins. When QCD events do enter the search region it is usually because one of the leading jets has been so severely mis-measured that it is not reconstructed as one of the leading jets.

A region where $\not{E}_T$ is closely aligned to one of the leading jets is ideal for

the QCD control region. This is achieved by removing the QCD killing cuts ($\Delta\phi_{1234}$) and instead requiring that $\Delta\phi$ between $\not\!\!E_{\mathrm{T}}$ and one of the three leading jets is below 0.1 ($\Delta\phi_{123} < 0.1$). The search region binning is then applied to this control region. As with the other backgrounds, the $\not\!\!E_{\mathrm{T}}$ shape is not dependent on N_{b} within the current statistics so these bins are combined in order to reduce the statistical uncertainties.

Very few generated events remain after the baseline selection in the currently available QCD MC samples. Some of these events have weights on the order of one event when scaled to 2.3 fb^{-1}. Therefore the effective luminosity of the sample is first increased with a method called local smearing.

Local Smearing of QCD MC

Introducing events that have different mis-measurements of the leading jets is a feasible way to increase the effective sample size without the enormous computational cost of producing significantly more QCD events.

The mis-measurement of a jet can be parameterized by the jet response. This is defined as:

$$r_{\mathrm{jet}} = \frac{p_{\mathrm{T}}(\mathrm{reco})}{p_{\mathrm{T}}(\mathrm{gen})}$$

where p_{T}(reco) and p_{T}(gen) are a jet's reconstructed and generator level p_{T} respectively. The r_{jet} distribution is constructed using the two highest generator level p_{T} jets in QCD MC events. It is binned in both generator level p_{T} and b vs non-b jets to account for differences from jet p_{T} and flavor.

A set of smeared events is created for every event in the original sample by assigning new values of r_{jet} to each of the two highest p_{T} jets in the event. The new values of r_{jet} are randomly taken from a window around the jet's original r_{jet} value

using the distribution from the appropriate p_{T}-flavor bin. The window size is dependent on the original value of r_{jet}. A tight window is used near $r_{\mathrm{jet}} = 1$ where the distribution is well populated and can change quite rapidly. A larger window is used for the smaller and larger values of r_{jet} in the tails of the distribution where the statistics are much more limited.

The new reconstructed jet p_{T} is then used to recalculate all jet-related event variables. Every event is smeared 100 times to create the new smeared QCD sample. Statistical uncertainties are estimated with bootstrapping [36], using 50 pseudo experiments.

Jet Response Corrections

One complication with this control region is that significantly more of the events in the search region are in the tails of the r_{jet} distribution (i.e. the mis-measurements are very severe). As such, the extrapolation to the search region relies on r_{jet} being correctly modeled in MC. Corrections to r_{jet} in MC are extrapolated from data to account for any mis-modeling effects.

Since the definition of r_{jet} relies on the knowledge of a jet's generator level p_{T}, an analogous quantity needs to be defined for data. This is done by noting that, for events in this control region, $\not{E}_{\mathrm{T}}$ can be assumed to be a result of energy being missed during the reconstruction of the jet nearest to it. Therefore an estimate of the actual p_{T} of this jet can be obtained by adding the $\not{E}_{\mathrm{T}}$ back into it. This quantity is defined as the pseudo-generator level p_{T}: $p_{\mathrm{T}}(\text{pseudo gen}) = p_{\mathrm{T}}(\text{reco}) + \not{E}_{\mathrm{T}}$. It is used to define a pseudo jet response:

$$r_{\text{pseudo jet}} = \frac{p_{\mathrm{T}}(\text{reco})}{p_{\mathrm{T}}(\text{pseudo gen})}.$$

Since this correction will eventually be applied primarily based on generator-level information from the most mis-measured jet, the cuts on N_{J} and N_{b} are removed in order to increase the statistics of its derivation. Five different regions are used in order to obtain correction factors which are sensitive to both jet flavor and r_{jet}. First, events are divided into two regions based on whether the jet most closely aligned to $\not{E}_{\mathrm{T}}$ is a medium b-jet or is not a loose b-jet. Then the medium b-jet region is sub-divided into $r_{\mathrm{pseudo\ jet}} < 0.5$ and > 0.5. The non-bjet region is sub-divided into $r_{\mathrm{pseudo\ jet}} < 0.33$, $0.33 - 0.66$, and > 0.66. In MC the most mis-measured jet in each event is assigned to a truth category based on whether or not it is a generator level b-jet and its r_{jet} value by using the same requirements as $r_{\mathrm{pseudo\ jet}}$ for b/non-b jets above.

The scale factors are then calculated with the following equations:

$$\begin{bmatrix} D_1 - O_1 \\ \vdots \\ D_5 - O_5 \end{bmatrix} = \begin{pmatrix} MC_{1,1} & \cdots & MC_{1,5} \\ \vdots & \ddots & \vdots \\ MC_{5,1} & \cdots & MC_{5,5} \end{pmatrix} \cdot \begin{bmatrix} SF_1 \\ \vdots \\ SF_5 \end{bmatrix},$$

where the $D_i - O_i$ are the data yields in each of the control regions after subtracting off the expected number of non-QCD events. The number of non-QCD events is estimated by applying the other background prediction methods to the QCD control region. The $MC_{i,j}$ are the QCD MC yields for each truth category in each control region. And the SF_i are the scale factors for each truth category. These scale factors range from 0.44-1.13 and are applied per event based on the jet most aligned to $\not{E}_{\mathrm{T}}$.

Background Prediction

After applying the r_{jet} corrections to the smeared QCD MC sample, the QCD prediction can be carried out it. The number of data events observed in each bin in the control region is translated to the corresponding bins in the search region by using a transfer factor which is obtained from MC. For each search region bin the transfer factor is defined as:

$$TF_{\text{QCD}} = \frac{N_{\text{MC}}(\Delta\phi > 0.5)}{N_{\text{MC}}(\Delta\phi < 0.1)}$$

where $N_{\text{MC}}(\Delta\phi > 0.5)$ and $N_{\text{MC}}(\Delta\phi < 0.1)$ are the number of events in QCD MC for a search region bin and its corresponding control region bin respectively. The QCD prediction is then given by:

$$N_{\text{pred}}^{\text{QCD}} = TF_{\text{QCD}} \cdot \left[N_{\text{data}}(\Delta\phi < 0.1) - N_{\text{MC}}^{\text{other}}(\Delta\phi < 0.1)\right]$$

where $N_{\text{data}}(\Delta\phi < 0.1)$ is the number of observed data events in a given control region bin and $N_{\text{MC}}^{\text{other}}(\Delta\phi < 0.1)$ is the estimated number of non-QCD events in that bin. Figure 3.4 shows the comparison between data and MC across all of the QCD control region bins. The QCD prediction and transfer factors for each bin of the search region is given in table 3.4. The number of data events and expected contamination from non-QCD events in the control region bins are also included.

Uncertainties

The dominant uncertainty is from the limited data and MC statistics with bins ranging from 7-100% and 23-100% respectively. The MC uncertainties tend to be dominant in the lower $\not{E}_{\text{T}}$ regions while the data uncertainties tend to be

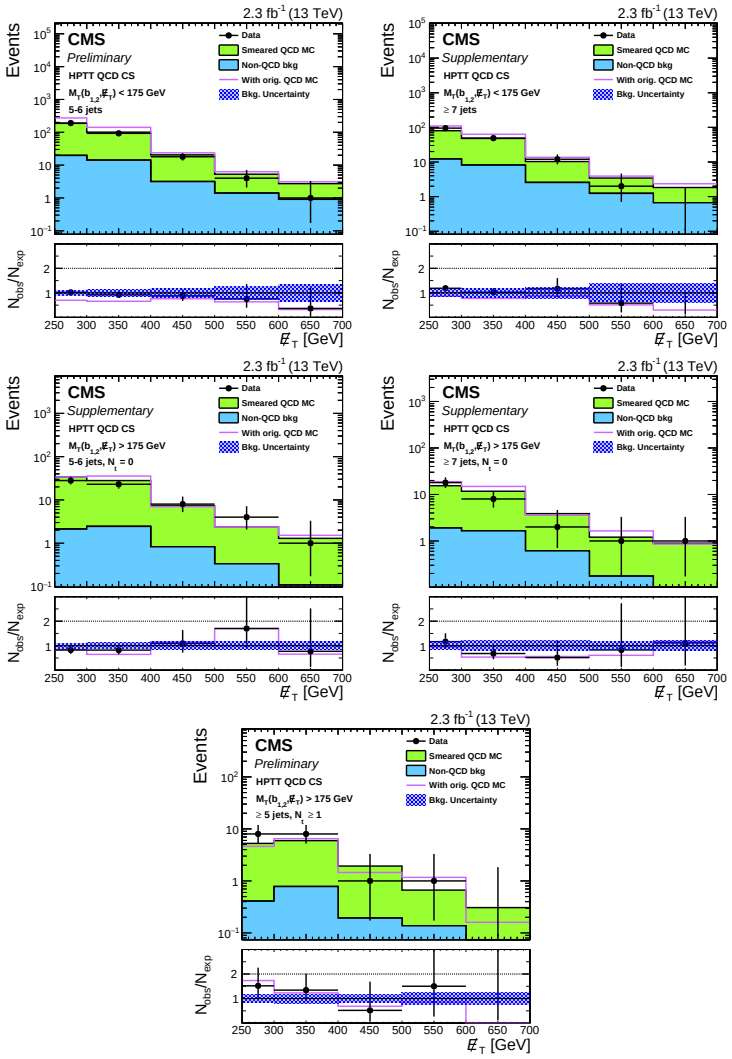

Figure 3.4: Data-MC comparisons across all bins in the QCD control region. The black points and stacked histograms are data and the contributing samples in MC respectively. In the data-over-MC ratio plot below each histogram, the black error bars are the statistical uncertainties on data and the blue hatched band are the statistical uncertainties on MC and the systematic uncertainty from the non-QCD background subtraction. The light purple line is the prediction using the original unsmeared MC without the r_{jet} correction. Each plot shows the $\not{E}_{\text{T}}$ bins for the given bin in $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}})$, N_{J}, and N_{t}: $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}}) < 175$, $N_{\text{J}} \in [5,6]$ (top left); $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}}) < 175$, $N_{\text{J}} \geq 7$ (top right); $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}}) \geq 175$, $N_{\text{J}} \in [5,6]$, $N_{\text{t}} = 0$ (middle left); $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}}) \geq 175$, $N_{\text{J}} \geq 7$, $N_{\text{t}} = 0$ (middle right); and $M_{\text{T}}(\text{b}_{1,2}, \not{E}_{\text{T}}) \geq 175$, $N_{\text{t}} \geq 1$ (bottom).

$\not\!\!E_{\mathrm{T}}$[GeV]	N_{data}	$N_{\mathrm{MC}}^{\text{non-QCD}}$	TF_{QCD}	$N_{\mathrm{pred}}^{\mathrm{QCD}}$	TF_{QCD}	$N_{\mathrm{pred}}^{\mathrm{QCD}}$
			$N_{\mathrm{b}} = 1$		$N_{\mathrm{b}} \geq 2$	
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) < 175$ GeV, $5-6$ jets						
250−300	185	18.86 ± 0.79	0.023 ± 0.007	3.773 ± 0.277 (stat.) ± 1.421 (syst.)	0.006 ± 0.002	1.072 ± 0.079 (stat.) ± 0.398 (syst.)
300−400	92	13.53 ± 0.73	0.018 ± 0.006	1.398 ± 0.146 (stat.) ± 0.586 (syst.)	0.006 ± 0.002	0.471 ± 0.049 (stat.) ± 0.208 (syst.)
400−500	16	3.02 ± 0.26	0.014 ± 0.008	0.188 ± 0.047 (stat.) ± 0.117 (syst.)	0.011 ± 0.007	0.144 ± 0.036 (stat.) ± 0.097 (syst.)
500−600	4	1.34 ± 0.17	0.049 ± 0.036	0.146 ± 0.073 (stat.) ± 0.117 (syst.)	0.016 ± 0.012	0.048 ± 0.024 (stat.) ± 0.043 (syst.)
>600	1	0.87 ± 0.21	0.015 ± 0.004	0.010 ± 0.010 (stat.) ± 0.005 (syst.)	0.001 ± 0.001	0.001 ± 0.001 (stat.) ± 0.001 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) < 175$ GeV, ≥ 7 jets						
250−300	93	11.58 ± 0.52	0.024 ± 0.005	1.918 ± 0.199 (stat.) ± 0.593 (syst.)	0.014 ± 0.005	1.167 ± 0.121 (stat.) ± 0.488 (syst.)
300−400	48	7.73 ± 0.43	0.028 ± 0.009	1.107 ± 0.160 (stat.) ± 0.446 (syst.)	0.009 ± 0.004	0.360 ± 0.052 (stat.) ± 0.176 (syst.)
400−500	15	2.44 ± 0.26	0.045 ± 0.021	0.563 ± 0.145 (stat.) ± 0.294 (syst.)	0.009 ± 0.004	0.108 ± 0.028 (stat.) ± 0.057 (syst.)
500−600	1	1.18 ± 0.27	0.037 ± 0.017	0.024 ± 0.024 (stat.) ± 0.014 (syst.)	0.004 ± 0.002	0.003 ± 0.003 (stat.) ± 0.002 (syst.)
>600	0	0.63 ± 0.15	0.033 ± 0.015	0.021 ± 0.021 (stat.) ± 0.013 (syst.)	0.012 ± 0.007	0.008 ± 0.008 (stat.) ± 0.006 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} = 0$, $5-6$ jets						
250−300	30	1.89 ± 0.25	0.026 ± 0.008	0.728 ± 0.133 (stat.) ± 0.363 (syst.)	0.017 ± 0.009	0.471 ± 0.086 (stat.) ± 0.255 (syst.)
300−400	22	2.18 ± 0.30	0.031 ± 0.011	0.608 ± 0.130 (stat.) ± 0.240 (syst.)	0.018 ± 0.009	0.362 ± 0.077 (stat.) ± 0.190 (syst.)
400−500	7	0.74 ± 0.14	0.075 ± 0.033	0.474 ± 0.179 (stat.) ± 0.246 (syst.)	0.026 ± 0.021	0.162 ± 0.061 (stat.) ± 0.133 (syst.)
500−600	4	0.30 ± 0.14	0.117 ± 0.067	0.409 ± 0.205 (stat.) ± 0.260 (syst.)	0.003 ± 0.002	0.009 ± 0.004 (stat.) ± 0.007 (syst.)
>600	1	0.10 ± 0.04	0.050 ± 0.031	0.046 ± 0.046 (stat.) ± 0.030 (syst.)	0.015 ± 0.007	0.014 ± 0.014 (stat.) ± 0.007 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} = 0$, ≥ 7 jets						
250−300	18	1.83 ± 0.35	0.075 ± 0.025	1.210 ± 0.285 (stat.) ± 0.455 (syst.)	0.030 ± 0.012	0.478 ± 0.113 (stat.) ± 0.210 (syst.)
300−400	8	1.59 ± 0.31	0.115 ± 0.045	0.792 ± 0.280 (stat.) ± 0.348 (syst.)	0.011 ± 0.006	0.076 ± 0.027 (stat.) ± 0.049 (syst.)
400−500	2	0.59 ± 0.12	0.047 ± 0.023	0.079 ± 0.056 (stat.) ± 0.043 (syst.)	0.004 ± 0.003	0.007 ± 0.005 (stat.) ± 0.005 (syst.)
500−600	1	0.17 ± 0.05	0.019 ± 0.014	0.016 ± 0.016 (stat.) ± 0.013 (syst.)	0.007 ± 0.005	0.006 ± 0.006 (stat.) ± 0.004 (syst.)
>600	1	0.07 ± 0.03	0.028 ± 0.013	0.025 ± 0.025 (stat.) ± 0.013 (syst.)	0.023 ± 0.012	0.021 ± 0.021 (stat.) ± 0.014 (syst.)
$M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ GeV, $N_{\mathrm{t}} \geq 1$, ≥ 5 jets						
250−300	8	0.37 ± 0.06	0.036 ± 0.023	0.270 ± 0.095 (stat.) ± 0.186 (syst.)	0.008 ± 0.005	0.056 ± 0.020 (stat.) ± 0.038 (syst.)
300−400	8	0.70 ± 0.17	0.003 ± 0.002	0.021 ± 0.007 (stat.) ± 0.012 (syst.)	0.005 ± 0.004	0.032 ± 0.011 (stat.) ± 0.030 (syst.)
400−500	1	0.17 ± 0.04	0.145 ± 0.112	0.132 ± 0.132 (stat.) ± 0.106 (syst.)	<0.001	-
500−600	1	0.12 ± 0.04	0.151 ± 0.114	0.122 ± 0.122 (stat.) ± 0.097 (syst.)	0.005 ± 0.005	0.004 ± 0.004 (stat.) ± 0.004 (syst.)
>600	0	0.03 ± 0.02	0.082 ± 0.053	0.075 ± 0.075 (stat.) ± 0.049 (syst.)	0.003 ± 0.002	0.003 ± 0.003 (stat.) ± 0.002 (syst.)

Table 3.4: The QCD prediction and transfer factors for each bin of the search region. The number of data events and expected contamination from non-QCD events in the control region bins are also included.

dominant in the higher $\not\!\!E_{\mathrm{T}}$ regions.

The uncertainty on the non-QCD background subtraction in the control region is taken to be the full size of the subtraction. It has a significant contribution in most bins, ranging between 7-35%. Uncertainties on the jet response tail correction can reach 15% and include the data/MC statistics, non-QCD subtraction, and effects from b-tagging. Standard corrections and systematic uncertainties for pileup reweighting (1-19%), jet energy scales (1-19%), b-tagging (0-8%), and the lepton (0-1%) and tau vetoes (0-7%) are also applied. They are discussed further in section 3.5.

3.4 ttZ

ttZ look almost identical to signal events when the Z decays to neutrinos and both tops decay hadronically since this results in large $\not{E}_{\mathrm{T}}$ and a significant numbers of jets. These events also pass the other selection and binning requirements such as N_{b} and $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$ since these jets are from real top decays. And since the $\not{E}_{\mathrm{T}}$ is usually from real neutrinos, the $\Delta\phi_{1234}$ cuts have a similar effect as they do on signal. This means that defining a control region that has both a high purity of ttZ events that are kinematically similar to those in the search region and low signal contamination is not feasible.

Luckily, the cross section for ttZ is very low so its contribution to the search region is quite small. However, it does becomes more significant in the higher $\not{E}_{\mathrm{T}}$ bins, especially those with one top. Unlike the QCD, the available ttZ MC sample has sufficient statistics to provide reasonable estimates and there are no complications from rare effects. Thus the ttZ background prediction is taken directly from the MC.

Uncertainties

The dominant uncertainty in most bins is from the limited MC statistics, ranging from 13-100%. Based on the 8 TeV CMS measurement [37], a 30% uncertainty is assigned for the cross section. Since this sample is normalized directly to the collected luminosity rather than to a data control region, the 2.7% uncertainty on the integrated luminosity measurement is taken as a systematic uncertainty. Other theoretical uncertainties that at least partially cancel out between the search and control regions in the data-driven background estimations need to be included for the ttZ prediction. These include renormalization scales and

PDFs and range from a few percent to around 25%. An uncertainty of 10% is assigned to account for data-MC differences in the top-tagging efficiency in the one top bins (section 3.5.5). Standard corrections and systematic uncertainties for pileup reweighting (0-13%), jet energy scales (0-20%), b-tagging (0-8%), and the lepton vetoes (0-5%) are also applied (section 3.5).

3.5 Corrections to Simulation

Uncertainties on the effects of various corrections are usually propagated to the final prediction by varying a given correction by its uncertainty and then re-running the prediction for all affected backgrounds. For each bin in the search region the percent change in the predicted number of events is taken as a systematic uncertainty on the prediction in that bin. These uncertainties are correlated between bins in the final prediction. For corrections that are relevant for multiple background/signal processes, separate systematic uncertainties are calculated for each process and then correlated between processes in the final prediction.

3.5.1 Pileup

As discussed in section 2.1.2, the pileup profile used for the production of the MC samples is not the same as what is observed in the data. To correct for this, MC events are reweighted such that they match the data. Systematic uncertainties are propagated to the prediction by varying the minimum bias cross section used in the reweighting by $\pm 5\%$. These corrections and uncertainties are applied to all MC samples.

3.5.2 Leptons

Scale factors are used to correct the MC for any observed differences in lepton selection efficiencies with respect to data. These are centrally derived by the SUSY Lepton Scale Factor Group using the Tag-and-Probe method. In this method, a tight selection is used to get a clean sample of $Z \to \ell\ell$ events where one "tag" lepton is well identified. The efficiency of identifying the second "probe" lepton is then measured. The scale factors are defined as the ratio of the efficiency in data over the efficiency in MC and measured separately for electrons and muons and parameterized by kinematic observables.

The identification scale factors are parameterized by the lepton's p_T and η to account for any kinematic correlations in lepton ID efficiencies. Isolation efficiencies are measured for probe leptons that pass the identification requirements. A local activity variable is used along with the lepton's p_T to parameterize the isolation corrections because the regions used in SUSY analyses have higher levels of hadronic activity than the $Z \to \ell\ell$ sample the scale factors are derived from. This local activity variable is defined as the Σp_T of all PF candidates in a ΔR annulus around the lepton that extends from the cone used to compute the lepton's isolation out to 0.4.

For every identified isolated lepton in an event the scale factors for identification and isolation ($SF_{\text{ID sel}}$ and $SF_{\text{Iso sel}}$ respectively) appropriate for the lepton's p_T, η, and activity are applied to the overall event weight for MC events. A scale factor for every lepton candidate that fails either the identification or isolation requirements is also applied to the overall MC event weight. These scale factors are given by:

$$SF_{\text{ID fail}} = \frac{1 - \varepsilon_{\text{ID sel}} \cdot SF_{\text{ID sel}}}{1 - \varepsilon_{\text{ID sel}}}$$

when the candidate fails the identification requirement and by:

$$SF_{\text{Iso fail}} = SF_{\text{ID sel}} \cdot \frac{1 - \varepsilon_{\text{Iso sel}} \cdot SF_{\text{Iso sel}}}{1 - \varepsilon_{\text{Iso sel}}}$$

when the candidate passes the identification requirement but fails the isolation requirement. $\varepsilon_{\text{ID sel}}$ is the efficiency in MC for a candidate to pass the identification criteria and $\varepsilon_{\text{Iso sel}}$ is the efficiency in MC for a candidate that passes the identification criteria to also pass the isolation criteria. Both efficiencies are evaluated with leptonic $\text{t}\bar{\text{t}}$/W+jets events that pass the baseline selection cuts other than the lepton vetoes.

The systematic and statistical uncertainties on the Tag-and-Probe scale factors are propagated to systematic uncertainties on all backgrounds and signal samples.

3.5.3 Tau Veto

Any differences in the tau veto between data and MC also need to be corrected for. This correction is measured in a control region (CR_τ) with the tau veto requirement inverted and the remaining baseline selection cuts applied. The electron and muon vetoes are left in tact in order to keep this measurement independent from the electron and muon corrections. The correction factor is defined as the efficiency for identifying a tau in data divided by the efficiency in MC. It is calculated as:

$$\text{Corr}_\tau = \frac{N_{\text{data}}(\text{CR}_\tau) - SF_{\text{norm}} \cdot N_{\text{MC}}^{0 \text{ gen } \tau}(\text{CR}_\tau)}{SF_{\text{norm}} \cdot N_{\text{MC}}^{\geq 1 \text{ gen } \tau}(\text{CR}_\tau)}$$

where SF_{norm} is the normalization of MC to data in the control region before requiring a tau candidate. The MC is divided into two pieces: one with a gen-

erator level tau and the other without which are denoted as $N_{\mathrm{MC}}^{0\ \mathrm{gen}\ \tau}(\mathrm{CR}_\tau)$ and $N_{\mathrm{MC}}^{\geq 1\ \mathrm{gen}\ \tau}(\mathrm{CR}_\tau)$ respectively. Corr_τ is measured separately for candidates with $p_{\mathrm{T}} < 20$ and $p_{\mathrm{T}} \geq 20$ GeV.

The number of events in the search region is adjusted as:

$$N_{\mathrm{MC}}^{\mathrm{corr}} = N_{\mathrm{MC}} + (1 - \mathrm{Corr}_\tau) \cdot N_{\mathrm{MC}}^{\geq 1\ \mathrm{gen}\ \tau}(\mathrm{CR}_\tau).$$

The statistical uncertainties are propagated to the prediction as systematic uncertainties for the LLB and QCD background predictions.

3.5.4 b-tagging

The b-tagging scale factors are provided centrally by the CMS B-tag Physics Object Group and applied based on jet kinematics and flavor. The corrections for heavy flavor b and c-quarks are varied separately from the corrections light quark and gluon mis-tagging when propagating systematic uncertainties to the prediction of all backgrounds and expected signal yields.

3.5.5 Top Tagging

A semi-leptonic $\mathrm{t\bar{t}}$ control region is used to study the top-tagging efficiency. Events are required to pass a single muon trigger and have at least one b-jet that is in the same hemisphere as the muon in order to select for semi-leptonic tops. $\not{E}_{\mathrm{T}} > 50$ is also required in order to reduce QCD contamination. The efficiency is studied for top candidates in the opposite hemisphere as the muon. Since the efficiency in MC and data agree within 10%, this is assigned as a systematic uncertainty on $\mathrm{t\bar{t}}$, ttZ, and signal samples in the bins with one top.

A sample consisting primarily of light quarks and gluons is used to study the mis-tag rate. Events in this sample are required to have $\not{E}_{\mathrm{T}} > 200$ GeV and a high overall event energy ($H_{\mathrm{T}} > 1$ TeV where $H_{\mathrm{T}} = \Sigma p_{\mathrm{T}}$ for all jets in the event). The mis-tag rate is about 30% higher in MC so a correction is applied and a 10% systematic uncertainty is assigned to the W+jets, $\mathrm{Z} \to \nu\nu$, and QCD predictions in the one top bins.

3.6 Validation in Data Control Regions

The entire background estimation is carried out in control regions before unblinding the search region. These regions are chosen such that they have negligible signal contamination and do not overlap with the search region or any control regions. Three regions are used for this purpose: a zero b-tag region, a low $\not{E}_{\mathrm{T}}$ region, and a low jet region. All other baseline cuts and search region bins are applied except as noted below. Good agreement between data and MC is observed in all three control regions. In all of the following plots the black points and hatched blue bands are the statistical uncertainties for data and MC respectively. Three example signal points are shown to give a sense of the possible level of signal contamination.

The zero b-tag region replaces the two bins in N_{b} with a single bin: $N_{\mathrm{b}} = 0$. It is shown in figure 3.5. The bins in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$ are removed since by definition these events have no b-jets with which to calculate $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$.

The low $\not{E}_{\mathrm{T}}$ region, shown in figure 3.6, replaces the five $\not{E}_{\mathrm{T}}$ bins with a single $\not{E}_{\mathrm{T}}$ bin: $\not{E}_{\mathrm{T}} \in [200, 250)$.

The low jet region, shown in figure 3.7, replaces the bins of N_{J} with the single bin: $N_{\mathrm{J}} \in [2, 4]$. It also lowers the baseline $\not{E}_{\mathrm{T}}$ cut to 200 GeV to create two

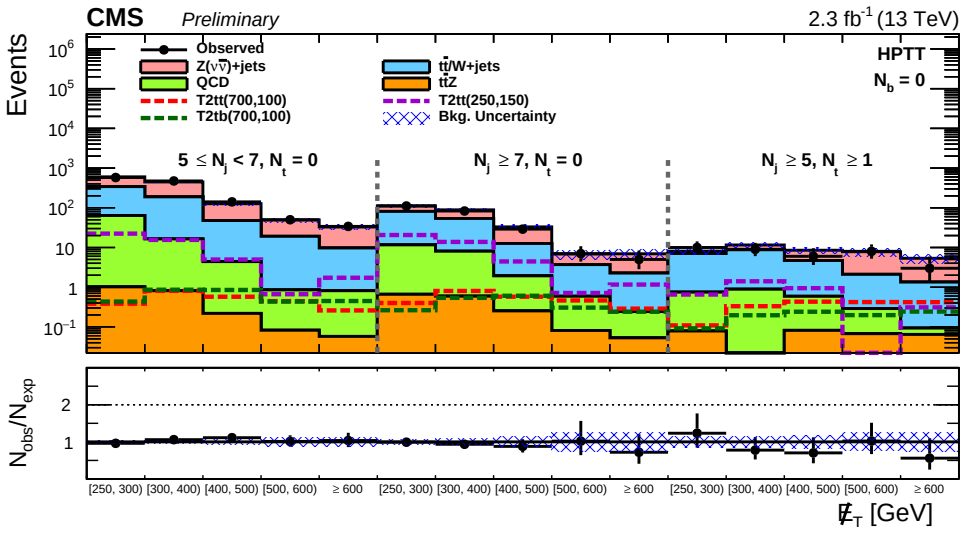

Figure 3.5: Validation of the full background prediction method in the $N_{\mathrm{b}} = 0$ region. The bins in N_{J}, N_{t}, and $\not{E}_{\mathrm{T}}$ are shown on the plots.

new $\not{E}_{\mathrm{T}}$ bins and merges the highest two $\not{E}_{\mathrm{T}}$ bins in the zero top regions. This final binning in $\not{E}_{\mathrm{T}}$ is: [200,225), [225,250), [250,300), [300,400), [400,500), and $\geq$ 500 GeV. The statistics are much more limited in the $N_{\mathrm{t}} \geq 1$ region so after lowering the baseline $\not{E}_{\mathrm{T}}$ cut only two $\not{E}_{\mathrm{T}}$ bins are used: [200,400) and $\geq$ 400 GeV.

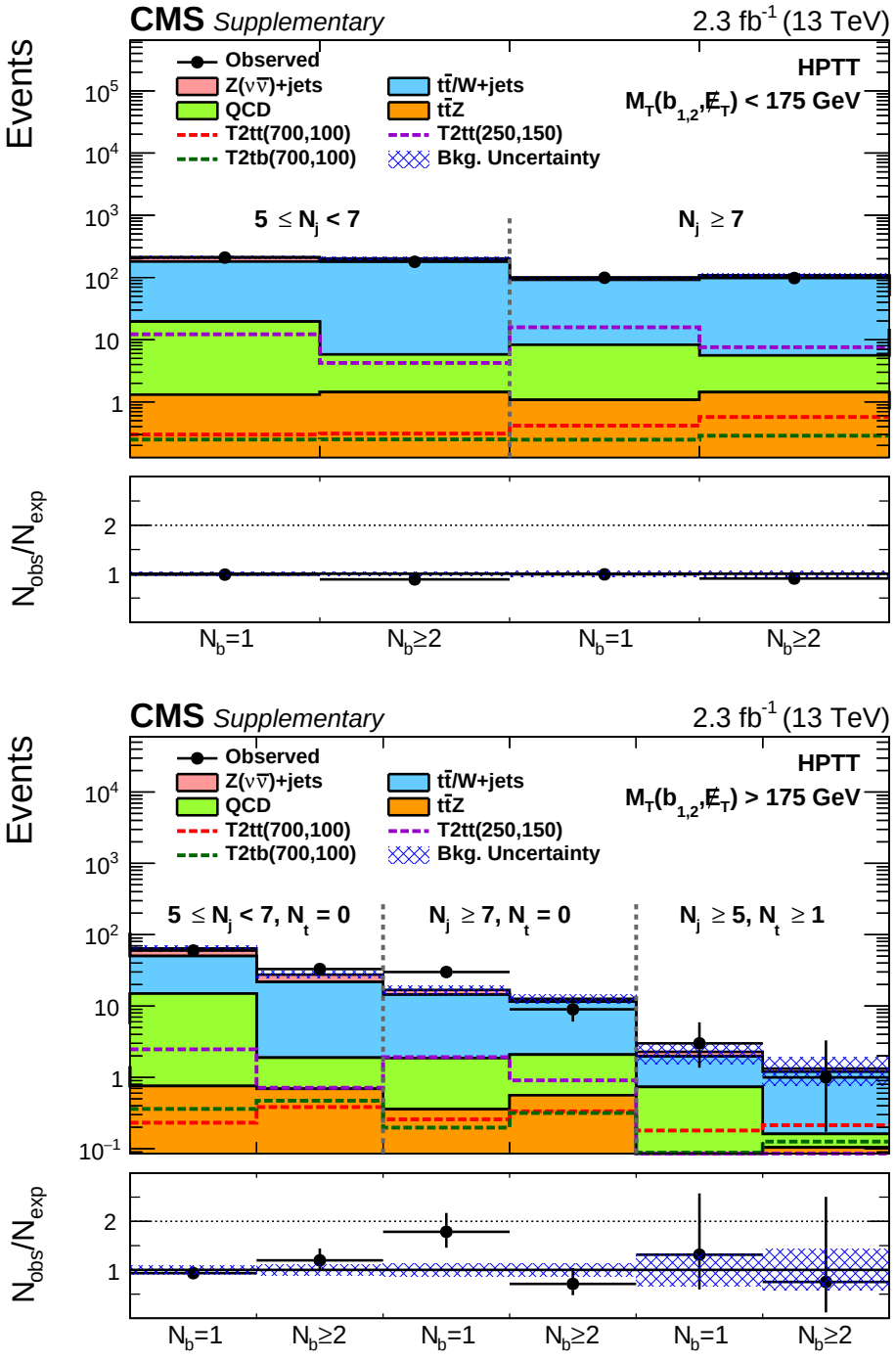

Figure 3.6: Validation of the full background prediction method in the low $\not\!\!E_{\mathrm{T}}$ region ($\not\!\!E_{\mathrm{T}} \in [200, 250)$). The regions with $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) < 175$ (top) and $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not\!\!E_{\mathrm{T}}) \geq 175$ (bottom) are shown. The additional bins in N_{J}, N_{t}, and N_{b} are shown on the plots.

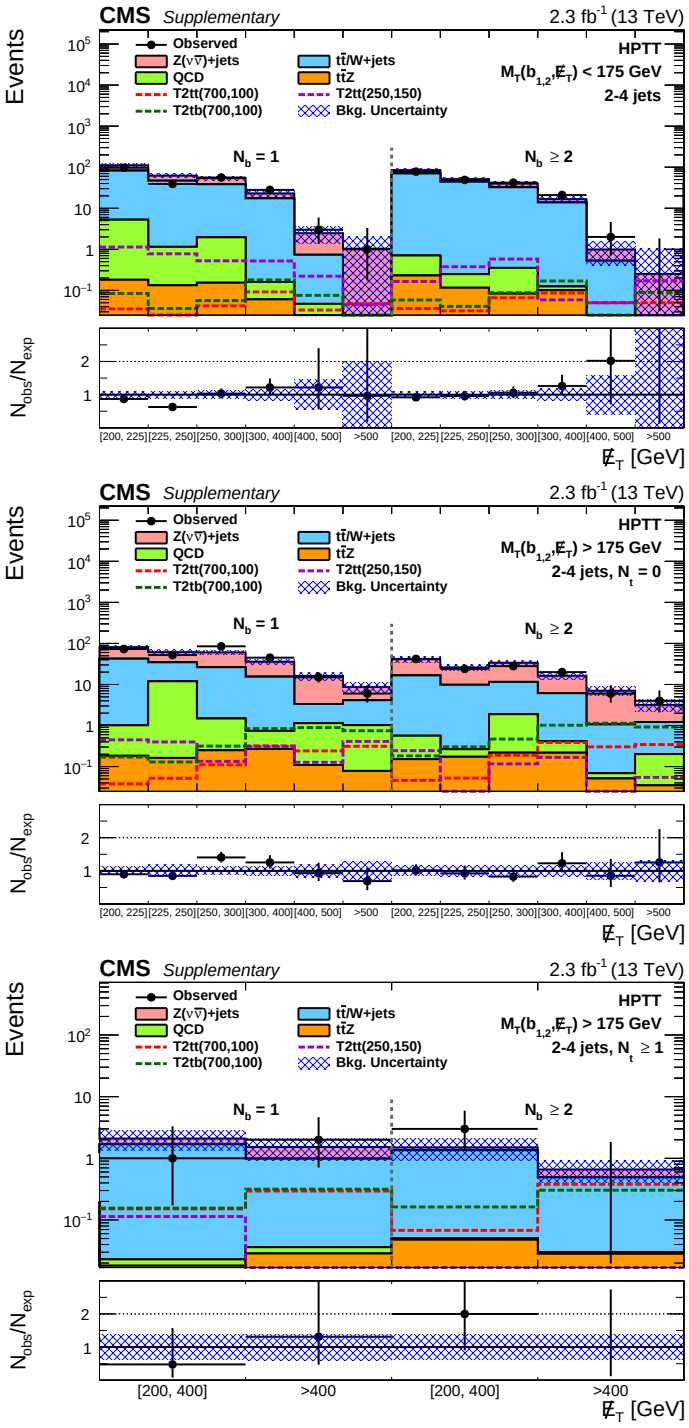

Figure 3.7: Validation of the full background prediction method in the low jet region ($N_{\mathrm{J}} \in [2, 4]$). From top to bottom, the plots show the regions with $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$; $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} = 0$; and $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} \geq 1$. The additional bins in N_{b} and $\not{E}_{\mathrm{T}}$ are shown on the plots.

Chapter 4

Results

The predicted SM background yields and observed events in data for each bin of the search region are shown in figure 4.1 and tabulated in table 4.1. The expected yields for a few example signal points are also shown in the plots. No statistically significant deviations from the SM background prediction are observed.

4.1 Interpretation

These results are interpreted in terms of the simplified models discussed in section 1.1.3. Exclusion limits are set for both the T2tt and T2tb simplified models since no statistically significant deviations from the prediction are observed in data. This is done by comparing the relative strengths of the hypothesis that the observed data only consists of background events (background-only) to the hypothesis that the observed data consists of both signal and background events (signal+background). The uncertainties on both the predicted backgrounds, as discussed throughout the previous chapter, and the expected signal yields, discussed below, are an important factor in determining how closely these hypotheses

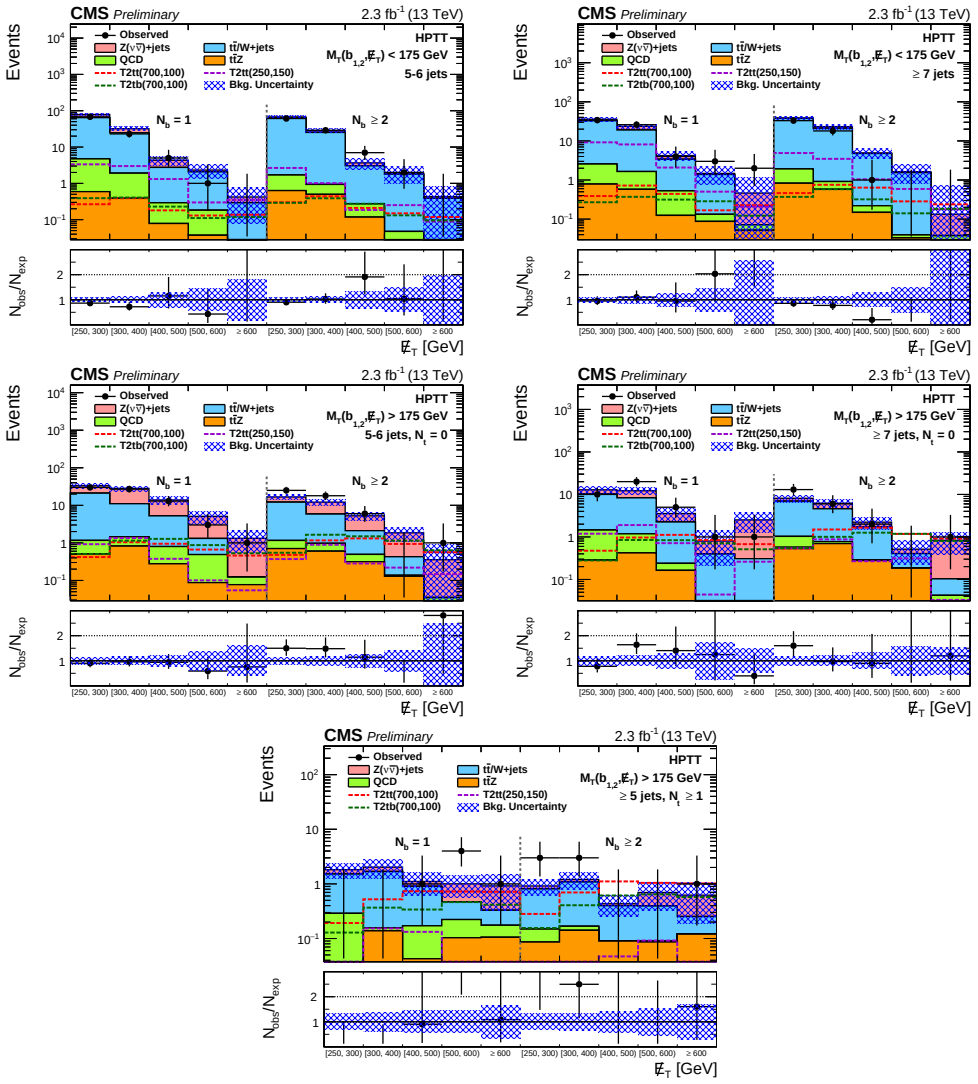

Figure 4.1: The predicted SM background yields and the number of events observed in data for every bin of the search region. The expected yields from several example signal points are also included as dashed lines. The error bars and blue hatched area are the statistical uncertainties on the data and MC respectively. Each plot shows the N_{b} and $\not{E}_{\mathrm{T}}$ bins for the given bin in $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}})$, N_{J}, and N_{t}: $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \in [5, 6]$ (top left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) < 175$, $N_{\mathrm{J}} \geq 7$ (top right); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \in [5, 6]$, $N_{\mathrm{t}} = 0$ (middle left); $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{J}} \geq 7$, $N_{\mathrm{t}} = 0$ (middle right); and $M_{\mathrm{T}}(\mathrm{b}_{1,2}, \not{E}_{\mathrm{T}}) \geq 175$, $N_{\mathrm{t}} \geq 1$ (bottom).

$\not\!\!E_T$[GeV]	$t\bar{t}$/W+jets	$Z \to \nu\nu$	QCD	ttZ	Total SM	Data
$M_T(b_{1,2}, \not\!\!E_T) < 175$ GeV, $5 \leq N_J < 7$, $N_b = 1$						
250-300	60 ± 6	14 ± 3	4.1 ± 1.7	0.59 ± 0.21	79 ± 7	68
300-400	23 ± 3	7.4 ± 1.9	1.5 ± 0.8	0.39 ± 0.14	32 ± 4	23
400-500	2.5 ± 1.0	1.6 ± 0.8	0.21 ± 0.15	0.08 ± 0.04	4.3 ± 1.3	5
500-600	1.9 ± 1.0	0.25 ± 0.27	0.14 ± 0.15	0.04 ± 0.02	2.3 ± 1.0	1
>600	0.28 ± 0.31	0.13 ± 0.15	0.01 ± 0.01	<0.01	0.42 ± 0.34	0
$M_T(b_{1,2}, \not\!\!E_T) < 175$ GeV, $5 \leq N_J < 7$, $N_b \geq 2$						
250-300	61 ± 6	4.7 ± 1.4	1.1 ± 0.5	0.63 ± 0.22	68 ± 6	61
300-400	24 ± 3	3.0 ± 1.0	0.44 ± 0.23	0.50 ± 0.18	28 ± 4	29
400-500	2.8 ± 1.2	0.61 ± 0.33	0.16 ± 0.13	0.12 ± 0.06	3.7 ± 1.2	7
500-600	1.7 ± 0.9	0.13 ± 0.15	0.05 ± 0.06	<0.01	1.9 ± 0.9	2
>600	0.38 ± 0.41	0.04 ± 0.06	<0.01	0.01 ± 0.01	0.43 ± 0.41	0
$M_T(b_{1,2}, \not\!\!E_T) < 175$ GeV, $N_J \geq 7$, $N_b = 1$						
250-300	30 ± 4	3.0 ± 1.0	1.8 ± 0.6	0.79 ± 0.28	36 ± 4	34
300-400	17 ± 3	4.6 ± 1.6	1.1 ± 0.5	0.58 ± 0.21	24 ± 3	26
400-500	2.9 ± 0.9	0.82 ± 0.64	0.40 ± 0.27	0.12 ± 0.07	4.2 ± 1.1	4
500-600	1.3 ± 0.7	0.09 ± 0.11	0.05 ± 0.05	0.09 ± 0.05	1.5 ± 0.7	3
>600	<0.56	0.39 ± 0.46	0.02 ± 0.02	0.05 ± 0.03	0.46 ± 0.72	2
$M_T(b_{1,2}, \not\!\!E_T) < 175$ GeV, $N_J \geq 7$, $N_b \geq 2$						
250-300	36 ± 4	0.96 ± 0.38	1.1 ± 0.5	0.83 ± 0.30	38 ± 4	33
300-400	20 ± 3	2.1 ± 0.9	0.34 ± 0.19	0.58 ± 0.22	23 ± 3	18
400-500	4.5 ± 1.4	0.15 ± 0.13	0.07 ± 0.05	0.15 ± 0.07	4.9 ± 1.4	1
500-600	1.5 ± 0.8	0.09 ± 0.11	0.01 ± 0.01	0.03 ± 0.03	1.6 ± 0.8	0
>600	<0.59	0.10 ± 0.12	0.01 ± 0.01	0.03 ± 0.02	0.13 ± 0.60	0
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $5 \leq N_J < 7$, $N_t = 0$, $N_b = 1$						
250-300	20 ± 3	12 ± 3	0.66 ± 0.37	0.50 ± 0.19	33 ± 5	30
300-400	9.6 ± 2.3	17 ± 4	0.63 ± 0.32	0.82 ± 0.27	28 ± 4	27
400-500	4.4 ± 1.9	8.6 ± 2.6	0.52 ± 0.35	0.28 ± 0.12	14 ± 3	13
500-600	0.82 ± 0.63	3.8 ± 1.8	0.40 ± 0.35	0.09 ± 0.06	5.1 ± 1.9	3
>600	<0.4	1.2 ± 0.7	0.05 ± 0.05	0.08 ± 0.04	1.3 ± 0.8	1
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $5 \leq N_J < 7$, $N_t = 0$, $N_b \geq 2$						
250-300	11 ± 2	4.5 ± 1.4	0.45 ± 0.27	0.70 ± 0.24	17 ± 3	25
300-400	4.9 ± 1.2	6.3 ± 1.8	0.37 ± 0.23	0.60 ± 0.22	12 ± 2	18
400-500	1.6 ± 0.7	3.1 ± 1.1	0.18 ± 0.17	0.31 ± 0.12	5.3 ± 1.4	6
500-600	0.29 ± 0.24	1.4 ± 0.8	0.01 ± 0.01	0.13 ± 0.06	1.9 ± 0.8	0
>600	<0.49	0.32 ± 0.20	0.01 ± 0.02	0.02 ± 0.02	0.36 ± 0.53	1
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $N_J \geq 7$ $N_t = 0$, $N_b = 1$						
250-300	8.8 ± 1.9	2.5 ± 1.0	1.2 ± 0.6	0.29 ± 0.18	13 ± 2	10
300-400	7.1 ± 1.8	3.9 ± 1.5	0.76 ± 0.46	0.42 ± 0.18	12 ± 2	20
400-500	2.0 ± 0.8	1.3 ± 0.7	0.08 ± 0.07	0.16 ± 0.09	3.6 ± 1.1	5
500-600	0.38 ± 0.40	0.40 ± 0.43	0.02 ± 0.02	<0.01	0.80 ± 0.59	1
>600	0.28 ± 0.33	2.2 ± 1.2	0.02 ± 0.03	<0.01	2.5 ± 1.2	1
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $N_J \geq 7$ $N_t = 0$, $N_b \geq 2$						
250-300	5.9 ± 1.3	1.2 ± 0.5	0.46 ± 0.24	0.57 ± 0.21	8.1 ± 1.5	13
300-400	3.8 ± 1.0	1.6 ± 0.7	0.08 ± 0.06	0.70 ± 0.26	6.2 ± 1.2	6
400-500	1.5 ± 0.6	0.48 ± 0.27	0.01 ± 0.01	0.28 ± 0.12	2.2 ± 0.7	2
500-600	0.22 ± 0.25	0.11 ± 0.12	0.01 ± 0.01	0.18 ± 0.08	0.51 ± 0.29	0
>600	0.06 ± 0.07	0.73 ± 0.44	0.02 ± 0.03	0.02 ± 0.03	0.84 ± 0.45	1
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $N_J \geq 5$, $N_t \geq 1$, $N_b = 1$						
250-300	1.2 ± 0.5	0.30 ± 0.25	0.26 ± 0.21	0.02 ± 0.03	1.8 ± 0.6	0
300-400	1.5 ± 0.7	0.34 ± 0.26	0.02 ± 0.01	0.14 ± 0.06	2.0 ± 0.8	0
400-500	0.73 ± 0.40	0.20 ± 0.22	0.13 ± 0.17	0.04 ± 0.05	1.1 ± 0.5	1
500-600	0.25 ± 0.22	0.54 ± 0.34	0.12 ± 0.16	0.10 ± 0.06	1.0 ± 0.4	4
>600	0.15 ± 0.33	0.59 ± 0.49	0.07 ± 0.07	0.11 ± 0.05	0.92 ± 0.60	1
$M_T(b_{1,2}, \not\!\!E_T) \geq 175$ GeV, $N_J \geq 5$, $N_t \geq 1$, $N_b \geq 2$						
250-300	0.66 ± 0.26	0.11 ± 0.09	0.06 ± 0.05	0.09 ± 0.05	0.92 ± 0.29	3
300-400	0.92 ± 0.39	0.12 ± 0.10	0.03 ± 0.03	0.14 ± 0.08	1.2 ± 0.4	3
400-500	0.31 ± 0.17	0.03 ± 0.04	<0.01	0.09 ± 0.06	0.43 ± 0.18	0
500-600	0.30 ± 0.30	0.30 ± 0.21	<0.01	0.09 ± 0.04	0.70 ± 0.37	0
>600	0.13 ± 0.29	0.37 ± 0.32	<0.01	0.12 ± 0.05	0.62 ± 0.43	1

Table 4.1: Predicted yields for each background and the number of observed events in data for each bin of the search region. The uncertainties on the background predictions include statistical and systematic uncertainties.

match the data.

Signal Uncertainties

The dominant uncertainties on the MC signal samples are limited statistics, which range from 5-100%. A 2.7% uncertainty is placed on the luminosity since the signal samples are normalized directly to the total luminosity. An uncertainty of 10% is assigned to account for data-MC differences in the top-tagging efficiency in the one top bins (section 3.5.5). The uncertainty on the simulation of the hadronic recoil becomes relevant at higher system p_{T}. From studies of the p_{T} description of initial state radiation (ISR) jets in di-leptonic $\mathrm{t\bar{t}}$, a 15% uncertainty is assigned for $\widetilde{t\bar{t}}$ between 400 and 600 GeV. This increases to 30% for $\widetilde{t\bar{t}} > 600$ GeV. Effects of variations of the renormalization and factorization scales on the signal acceptance are applied as uncertainties of up to 18%. Standard corrections and systematic uncertainties for pileup reweighting (1-58%), jet energy scales (1-47%), b-tagging (0-17%), and the lepton vetoes (0-6%) are also applied (section 3.5).

Statistical Interpretation

The method used to set these limits is a modified frequentist method that has been agreed upon by both the ATLAS and CMS collaborations [38]. This method is protected against potential downward fluctuations in the observed number of background events which could otherwise allow the exclusions of signals a search is not actually sensitive to.

In this method a binned likelihood function is first constructed using the uncertainties as nuisance parameters. It is then fit to the observed data separately for both the background-only and signal+background hypotheses by finding the values of the nuisance parameters which best fit the data. A test statistic is de-

fined based on a profile likelihood ratio and its observed value is calculated for the given signal strength (μ). Next, probability distribution functions (*pdf*s) are constructed using the test statistic and appropriately fit values of the nuisance parameters for both the background-only and signal+background hypotheses. The *pdf*s are integrated over the test statistic from the observed value to infinity to obtain p-values: p_μ and $1 - p_b$ for the signal+background and background-only hypotheses respectively. Finally, the confidence level is defined as:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b}.$$

Thus the signal is excluded at the (1-CL_s) confidence level. To find the lowest signal strength that can be excluded at 95% confidence level, μ is adjusted until $CL_s = 0.05$.

The above procedure is applied to every signal point in the $m_{\tilde{t}}$-$m_{\tilde{\chi}_1^0}$ plane for both the T2tt and T2tb models. The highest signal strength (i.e. cross section) which is not excluded at the 95% confidence level is determined for that point. Signal models are excluded when their NLO+NLL theoretical cross section [39] is higher the 95% upper limit on the production cross section.

Statistical uncertainties on MC samples are uncorrelated across all bins. Statistical uncertainties from data in control regions are correlated across bins that share a given control region. The systematic uncertainties are correlated across all bins and all relevant background and signal samples. All 50 search region and 21 single-lepton control region bins are fit to the observed data simultaneously to take any potential signal contamination in this control region into account.

Figures 4.2 and 4.3 show the expected and observed 95% CL exclusion limits for the T2tt and T2tb simplified models respectively. The band along the $m_{\tilde{t}} =$

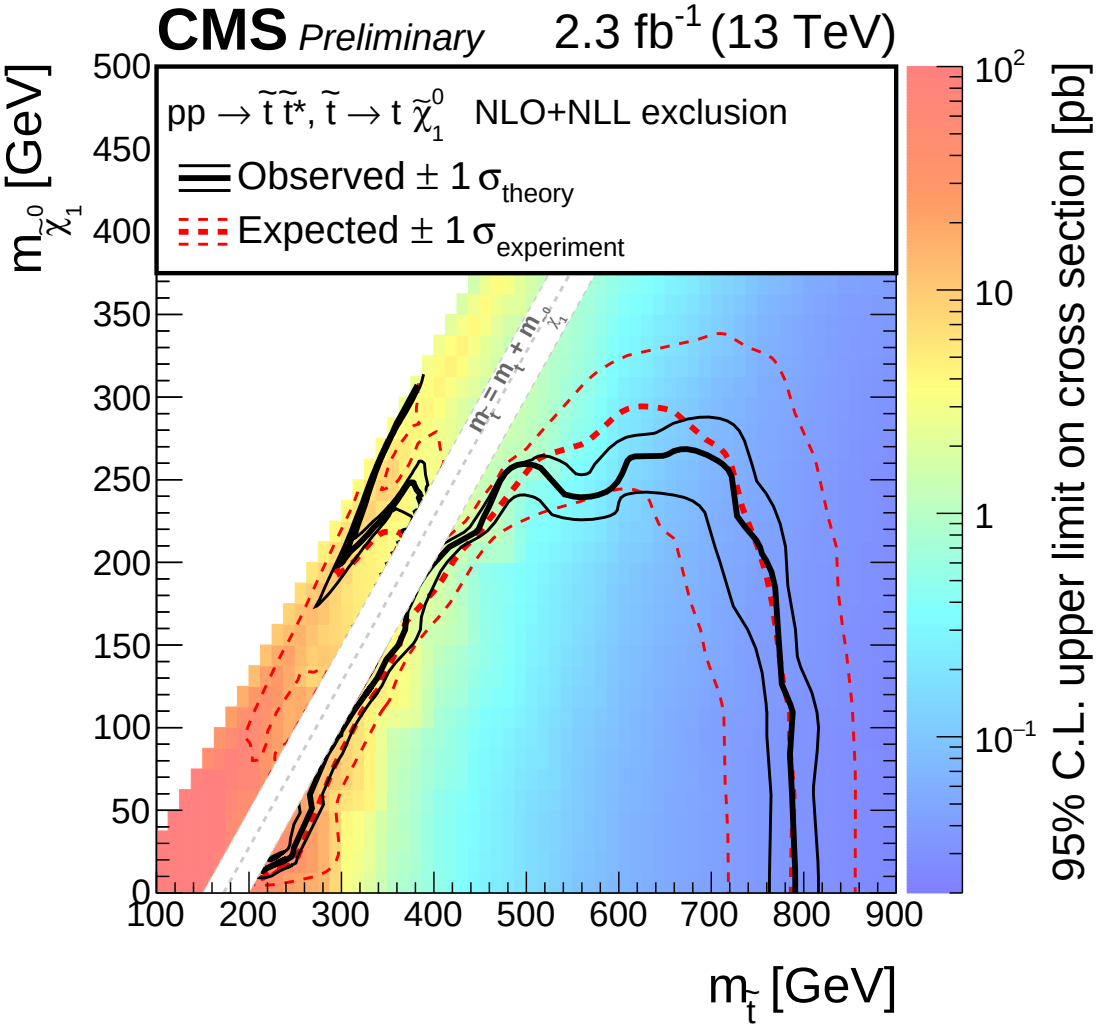

Figure 4.2: 95% CL exclusion limits for stop production in the T2tt model. The solid black and dashed red curves are the observed and expected exclusions contours (thick lines) respectively, with their corresponding ±1 standard deviations (thin lines).

$m_t - m_{\tilde{\chi}_1^0}$ diagonal is not shown since it was not included when these results were made public. This region is quite sensitive to signal contamination and systematic effects. Therefore a decision was made by the CMS SUSY group to not show it until after a careful and consistent treatment of these effects is agreed upon and implemented by all CMS stop searches.

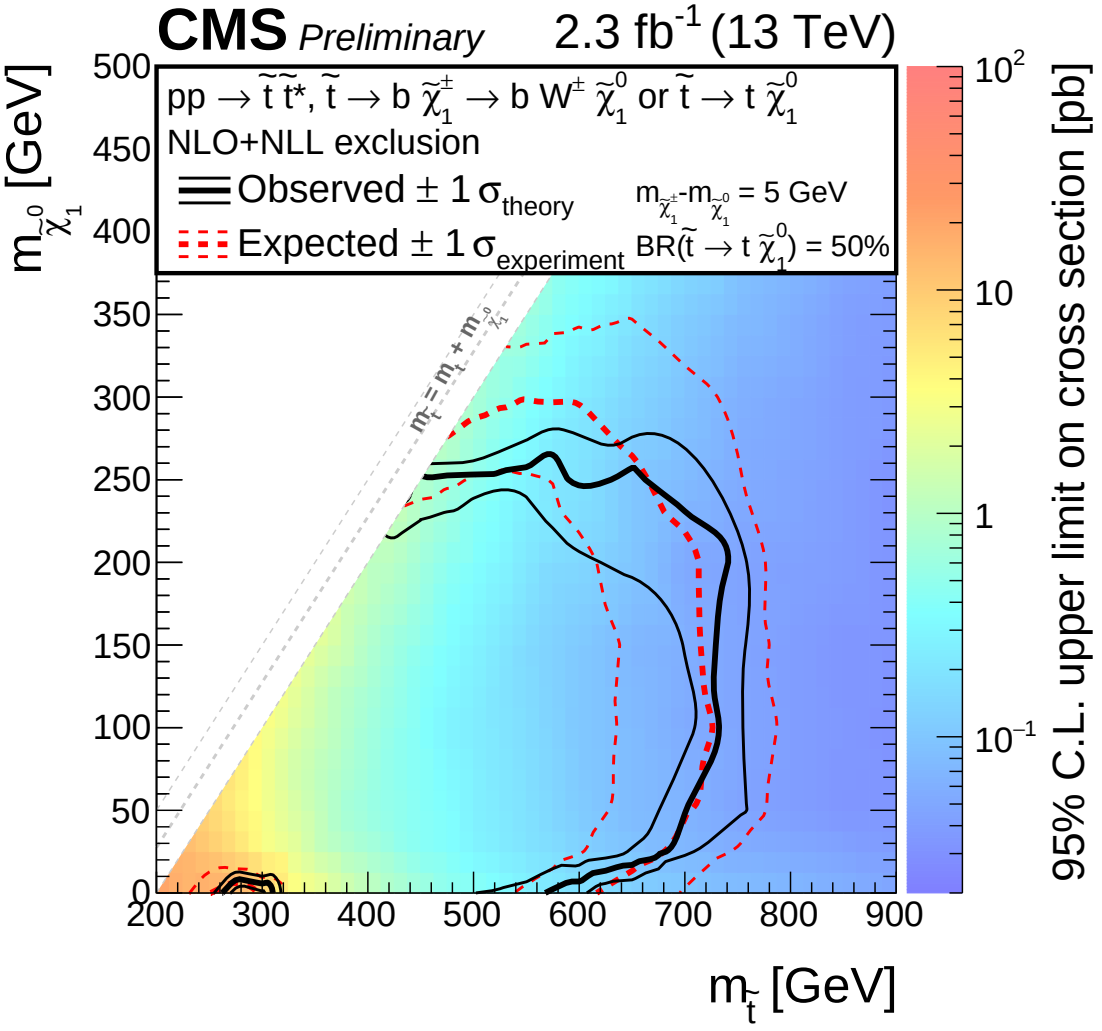

Figure 4.3: 95% CL exclusion limits for stop production in the T2tb model. The solid black and dashed red curves are the observed and expected exclusions contours (thick lines) respectively, with their corresponding ±1 standard deviations (thin lines).

4.2 Summary and Outlook

The analysis presented here has focused on creating a simple yet powerful search for stop pair production using the first data from Run 2. In particular, it is sensitive to three different decay modes and a variety of sparticle masses for

each. Compared to the data collected at 8 TeV during Run 1, the center-of-mass energy has increased by over 50%. Although the amount of data collected so far at 13 TeV is considerably lower, the sensitivity of this search is still comparable to 8 TeV searches. This improvement comes from the increased production cross section for stop pairs. Thus as more data is collected at 13 TeV future iterations of this search will be a powerful means by which SUSY could either be discovered or forced into smaller corners of phase space. There are also multiple fronts on which improvements can be made to this analysis in the future.

Short term, a procedure for handling the signal contamination and systematic effects in the corridor along the $m_{\tilde{t}} = m_t - m_{\tilde{\chi}_1^0}$ diagonal needs to be agreed upon and implemented across all the stop searches so that results in the entire probed $m_{\tilde{t}} - m_{\tilde{\chi}_1^0}$ plane can be shown. This search will also be combined with the stop search in the single lepton channel in the near future. A combination will take advantage of the orthogonality of the two searches and thus obtain more powerful exclusion limits. Although the treatment of correlated systematic uncertainties and any overlapping regions must be handled carefully, this search has been designed with a combination in mind. Exclusion limits for the case where both stops decay via an intermediate chargino can also be set once the full $m_{\tilde{t}} - m_{\tilde{\chi}_1^0}$ plane has been fully produced.

Moving forward into the rest of Run 2, there are several significant ways in which this analysis can be improved. The most obvious is the addition of more data as the increase in statistics will reduce the largest uncertainties. If there are enough statistics to sufficiently populate all of the bins in the various control regions the need for combining bins could be eliminated. This would completely eliminate the associated uncertainties. The addition of a W-tagger to compliment the top-tagger in the case where a stop decays via a chargino could increase the

sensitivity of the search to the T2tb and T2bW decay modes. At 8 TeV, jet substructure information was used to more efficiently reconstruct top-quark decays. This top-tagger could be resurrected and retrained. Some of the other variables used in analyses at 8 TeV could also be reevaluated for potential discriminating power in this analysis. Since many of them are likely to bring only modest discrimination, the possibility of combining various ones using a multivariate analysis such as the BDT could be pursued.

Bibliography

[1] **CMS** Collaboration, *Search for direct production of top squark pairs decaying to all-hadronic final states in pp collisions at sqrt(s) = 13 TeV*, Tech. Rep. CMS-PAS-SUS-16-007, CERN, Geneva, 2016. `http://cds.cern.ch/record/2141543`.

[2] D. Griffiths, *Introduction to Elementary Particles*. Wiley, second revised ed., 2008.

[3] C. G. Tully, *Elementary Particle Physics in a Nutshell*. Princeton University Press, 2011.

[4] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, Inc., 1984.

[5] **Particle Data Group** Collaboration, K. O. et al., *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001. 2015 update.

[6] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012) 30–61, [arXiv:1207.7235].

[7] **ATLAS** Collaboration, G. Aad *et. al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29, [arXiv:1207.7214].

[8] P. Langacker, *The Standard Model and Beyond*. CRC Press, 2010.

[9] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356. [Adv. Ser. Direct. High Energy Phys.18,1(1998)].

[10] "http://www.thphys.uni-heidelberg.de/ plehn/index.php?show=prospino." `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[11] J. Alwall, P. Schuster, and N. Toro, *Simplified Models for a First Characterization of New Physics at the LHC*, *Phys. Rev.* **D79** (2009) 075020, [arXiv:0810.3921].

[12] J. Alwall, M.-P. Le, M. Lisanti, and J. G. Wacker, *Model-Independent Jets plus Missing Energy Searches*, *Phys. Rev.* **D79** (2009) 015005, [arXiv:0809.3264].

[13] **LHC New Physics Working Group** Collaboration, D. Alves, *Simplified Models for LHC New Physics Searches*, *J. Phys.* **G39** (2012) 105005, [arXiv:1105.2838].

[14] L. Evans and P. Bryant, *LHC Machine*, *JINST* **3** (2008) S08001.

[15] C. Lefvre, "The CERN accelerator complex. Complexe des acclrateurs du CERN." `https://cds.cern.ch/record/1260465`, Dec, 2008.

[16] "CMS Integrated Luminosity, pp, 2015, $\sqrt{}$(s) = 13 TeV." `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`, Nov, 2015.

[17] **CMS** Collaboration, S. Chatrchyan *et. al.*, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008) S08004.

[18] M. Brice, M. Hoch, and J. Gobin, "View of the CMS Detector before closure." `https://cds.cern.ch/record/1133594`, Aug, 2008.

[19] D. Barney, "CMS slice." `https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=5581`, Sep, 2011.

[20] **CMS** Collaboration, *Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET*, Tech. Rep. CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr, 2009. `https://cds.cern.ch/record/1194487"`.

[21] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, *JHEP* **04** (2008) 063, [arXiv:0802.1189]. `dx.doi.org/10.1088/1126-6708/2008/04/063`, `arXiv:0802.1189`.

[22] **CMS** Collaboration, S. Chatrchyan *et. al.*, *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [arXiv:1211.4462]. `dx.doi.org/10.1088/1748-0221/8/04/P04013`, `arXiv:1211.4462`.

[23] **CMS** Collaboration, C. Collaboration, *Identification of b quark jets at the CMS Experiment in the LHC Run 2*, .

[24] **CMS** Collaboration, *Boosted Top Jet Tagging at CMS*, Tech. Rep. CMS-PAS-JME-13-007, CERN, Geneva, 2014. `http://cds.cern.ch/record/1647419`.

[25] **CMS** Collaboration, *Top Tagging with New Approaches*, Tech. Rep. CMS-PAS-JME-15-002, CERN, Geneva, 2016. https://cds.cern.ch/record/2126325.

[26] J. A. et al, *Madgraph 5: going beyond*, *Journal of High Energy Physics* **2011** (2011), no. 6 1–40. dx.doi.org/10.1007/JHEP06(2011)128, arXiv:hep-ph/0409146.

[27] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, [hep-ph/0409146]. dx.doi.org/10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[28] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, *JHEP* **11** (2007) 070, [arXiv:0709.2092]. dx.doi.org/10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[29] S. Alioli, P. Nason, C. Oleari, and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, *JHEP* **06** (2010) 043, [arXiv:1002.2581]. dx.doi.org/10.1007/JHEP06(2010)043, arXiv:1002.2581.

[30] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, *Eur. Phys. J.* **C71** (2011) 1547, [arXiv:1009.2450]. dx.doi.org/10.1140/epjc/s10052-011-1547-z, arXiv:0709.2092.

[31] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, *JHEP* **07** (2014) 079, [arXiv:1405.0301]. dx.doi.org/10.1007/JHEP07(2014)079, arXiv:1405.0301.

[32] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, *Comput. Phys. Commun.* **178** (2008) 852–867, [arXiv:0710.3820]. dx.doi.org/10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[33] **GEANT4** Collaboration, S. Agostinelli *et. al.*, *GEANT4: A Simulation toolkit*, *Nucl. Instrum. Meth.* **A506** (2003) 250–303. dx.doi.org/10.1016/S0168-9002(03)01368-8.

[34] **CMS** Collaboration, S. Abdullin, P. Azzi, F. Beaudette, P. Janot, and A. Perrotta, *The fast simulation of the CMS detector at LHC*, *J. Phys. Conf. Ser.* **331** (2011) 032049. 10.1088/1742-6596/331/3/032049.

[35] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, *PoS* **ACAT** (2007) 040, [physics/0703039].

[36] B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*. Society for Industrial and Applied Mathematics, 1982. http://epubs.siam.org/doi/abs/10.1137/1.9781611970319.

[37] **CMS** Collaboration, V. Khachatryan *et. al.*, *Observation of top quark pairs produced in association with a vector boson in pp collisions at* $\sqrt{s} = 8$ *TeV*, *JHEP* **01** (2016) 096, [arXiv:1510.0113]. dx.doi.org/10.1007/JHEP01(2016)096.

[38] **The ATLAS Collaboration, The CMS Collaboration, The LHC Higgs Combination Group** Collaboration, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011. https://cds.cern.ch/record/1379837.

[39] C. Borschensky, M. Krmer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, and X. Portell, *Squark and gluino production cross sections in pp collisions at* $\sqrt{s}$ *= 13, 14, 33 and 100 TeV*, *Eur. Phys. J.* **C74** (2014), no. 12 3174, [arXiv:1407.5066]. dx.doi.org/10.1140/epjc/s10052-014-3174-y, arXiv:1407.5066.