SEARCH FOR NEW PHYSICS IN SAME-SIGN DILEPTON EVENTS IN THE CMS DETECTOR AT THE LARGE HADRON COLLIDER

By

LANA MUNIZ

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2014

This is dedicated to my husband, Marty.

ACKNOWLEDGMENTS

I would like to thank my adviser John Yelton for providing helpful support and guidance during my time in graduate school. I also thank my committee members: Guenakh Mitselmakher, Konstantin Matchev, Andrey Korytov, David Tanner, and Vicki Sarajedini for their commitment to my education and interest in my research. I would like to thank the University of Florida faculty members and colleagues who participated in this CMS analysis effort through their contributions and oversight: Didar Dobur, Ronald Remington, Nikoloz Skhirtladze, Lesya Shchutska, Brandon Allen, Guenakh Mitselmakher, Andrey Korytov, and Konstantin Matchev. I would also like to acknowledge the hard work by the Tier-2 and HPC administrators, who provided key technical support for our research needs.

I would like to thank my family and friends for their loving support of my educational goals. I am very grateful to my father-in-law, Elliott McCrory, an accelerator physicist, for his encouragement over the years and for being a terrific role model. I want to thank Darsa Donelan, my bestie, for being my closest ally in graduate school and for an enduring friendship. I would especially like to acknowledge my husband, Marty. He has given me enormous support during graduate school, by moving his life to Florida and by his understanding during extended separations. Then, after our daughter, Emmy, was born in 2012, he was a very committed and supportive partner. It has been my immense privilege to share this journey with him.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure page

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

SEARCH FOR NEW PHYSICS IN SAME-SIGN DILEPTON EVENTS
IN THE CMS DETECTOR AT THE LARGE HADRON COLLIDER

By

Lana Muniz

August 2014

Chair: John M. Yelton
Major: Physics

CMS results of the new physics search in same-sign dilepton events with b-tagged jets and missing transverse energy, $\not\mathrm{E}_\mathrm{T}$, are presented. These results cover the full 2012 dataset from the LHC at $\sqrt{\mathrm{s}} = 8$ TeV, corresponding to an integrated luminosity of 19.5 fb^{-1}. Isolated same-sign dilepton events are comparatively easy to detect efficiently. They are predicted to be produced in abundance in some supersymmetry models, but are rarely produced in the Standard Model (SM) processes. Hence, this channel provides a very clean, low background, search for new physics. Multiple search regions defined by the observables $\not\mathrm{E}_\mathrm{T}$, hadronic energy ($\mathrm{H}_\mathrm{T}$), and number of b-tagged jets are considered. The yield of events in the data agrees with the SM prediction, therefore exclusion limits at 95% C.L. are presented for various simplified SUSY models.

CHAPTER 1
INTRODUCTION

In 2012, the long-anticipated discovery of the Higgs Boson was announced by scientists at the Large Hadron Collider (LHC) with a mass of 125 GeV. The Higgs Boson gives particles mass by the mechanism of a scalar field, and it was long-thought to be a missing piece of the Standard Model of Particle Physics (SM). It took two decades to build and commission the Compact Muon Solenoid (CMS) and ATLAS experiments, the two general-purpose detectors which independently discovered the Higgs. Scientists from dozens of participating countries contributed to this massive data analysis project. It has been a very exciting time in particle physics to witness this discovery unfold, providing us with a more complete picture of nature. The Higgs discovery, however, is just the end of a chapter in the story of particle physics.

Apart from the Higgs, the highest priority at the LHC for new physics to search for is supersymmetry (SUSY). SUSY models have particular appeal because they can specifically address remaining problems in the Higgs theory, notably the Hierarchy Problem (see Section 2.4). So far SUSY has been excluded only in the low mass regime, i.e. below the TeV-scale. SUSY would allow so many new particles and interactions in nature that, in principle, searches for new physics in many channels can be motivated by different viable versions of SUSY. However, new physics is most likely to be visible when the event signature is sparsely populated by Standard Model processes. Search strategies in such channels ought to remain general enough to catch an excess not arising from SUSY, too. This dissertation describes a particular SUSY-motivated search carried out at CMS, that of an excess in same-sign dilepton events. This signature is known for having a very low background from the Standard Model and thus any signal present is expected to be very clean. This, together with the prediction that such events should be produced in a number of SUSY models, makes it a closely-watched analysis within particle physics research.

New particles are produced in pp collisions at the LHC. One can think of $\mathrm{E} = \mathrm{mc}^2$ as a recipe, where energy mc^2 is needed to produce a particle at rest with mass m. The center of mass energy of a collision, denoted $\sqrt{\mathrm{s}}$, is the energy available to produce new and unstable particles. In pp collisions, each proton carries half the $\sqrt{\mathrm{s}}$ energy, but it is constituent quarks and gluons that actually collide, each carrying about $1/6$ of the proton energy. Thus, in order to produce and discover a new, heavy, particle, the $\sqrt{\mathrm{s}}$ of the machine must be many times larger than the mass of the particle. The LHC is the highest energy collider to date, achieving $\sqrt{\mathrm{s}} =$ 8 TeV, though it was designed to operate at $\sqrt{\mathrm{s}} =$ 14 TeV. Massive new particles are not observed directly in the detector as they are too short-lived, but rather events are sought out which include its likely decay products. To search for new physics, each decay signature, or channel, is analyzed independently to search for an excess of events above the expected SM background.

Particle collisions are governed by quantum mechanics, and hence their outcomes are probabilistic. The probability of a particular outcome is determined by its cross-section, or σ. The production of massive particles such as the top quark or the Higgs have a very small cross-section of order picobarns, where $1\,\mathrm{pb} = 10^{-36}\mathrm{cm}^2$. The production of lighter quarks and gluons have a large cross-section of order millibarns, where $1\,\mathrm{mb} = 10^{-27}\mathrm{cm}^2$. So for each Higgs produced, billions of mundane collisions had to occur. Luminosity, L, is defined such that luminosity times cross-section gives the number of collisions per second. The LHC has design $\mathrm{L} = 10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ with a proton-proton collision of $\sigma = 100\,\mathrm{mb} = 10^{-25}\mathrm{cm}^2$ at $25\,\mathrm{ns}$ between proton bunch crossings. This gives a collision rate of about $1\,\mathrm{GHz}$, or one Higgs produced every second. As of this writing, the highest luminosity achieved by the LHC has been $\mathrm{L} = 5 \times 10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ with $50\,\mathrm{ns}$ bunch crossings, to produce one Higgs every four seconds.

Each decay channel is associated with a *branching ratio*, BR, such that the sum of the BR for all decay channels is 100%. Further, a particular search strategy has a

unique event selection that has an efficiency, $\epsilon < 100\%$ of events for that channel. Thus, Eq. 1–1 gives the number of events one can expect to observe in a given analysis.

$$\mathrm{N} = \mathrm{L}\sigma \times \mathrm{BR} \times \epsilon \tag{1–1}$$

In this thesis LHC data is analyzed for signals in same-sign dilepton channels for which $\sigma \times \mathrm{BR}$ is predicted by SUSY models. Efforts are made to make ϵ as large as possible, so that the product shown in Eq. 1–1 gives a calculated number of events that could, in principle, be observed. Great efforts are made to suppress the backgrounds so that even a modest number of produced signal events would constitute a significant observation. In fact, no significant excess was found, but using the known luminosity and efficiencies, this leads to constraints on the $\sigma \times \mathrm{BR}$ for SUSY production, and thus rules out some particular SUSY models.

The organization of this dissertation is as follows. Chapter 2 describes the Standard Model of particles and their interactions . Chapter 3 gives a summary of and motivation for SUSY and how it pertains to the physics of the same-sign dilepton channel. Chapter 4 gives an overview of the LHC. Chapter 5 describes the CMS detector. Chapter 6 describes a recent study on whether vertexing, among other strategies, could be used to increase the sensitivity of the same-sign dilepton analysis in future iterations. Chapter 7 is an in-depth analysis of LHC data, a search for new physics in the same-sign dilepton channel, including a parametrization of the analysis acceptance intended for use by those outside the CMS collaboration and limits set on specific SUSY models. Chapter 8 provides the conclusion to this new physics search.

CHAPTER 2
THE STANDARD MODEL OF PARTICLE PHYSICS

2.1 Particles and Interactions

The Standard Model (SM) of particle physics describes the fundamental particles of nature and the forces that govern their interactions. Fermions, consisting of quarks and leptons, constitute matter and have half-integer spin, while gauge bosons are force carriers and have spin 1. Photons carry the electromagnetic force, $W^{\pm}$ and Z^0 bosons mediate the weak force of nuclear decay, and gluons mediate the strong force which binds quarks into protons and neutrons. The latest addition to the SM is the scalar Higgs field, which permeates all of space and gives mass to fermions and some bosons. Its carrier is the Higgs boson, which has spin 0.

Leptons, shown in Table 2-1, come in three flavors, or generations: electrons, muons, and taus. Each generation contains a charged lepton and a neutral lepton, called a *neutrino*. Charged leptons are governed by both the weak and electromagnetic forces, and they have integer charge. Neutrinos only interact by the weak force since they have no electric charge. Within the last decade, the surprise discovery that neutrinos oscillate between flavors, and therefore have non-zero mass, has led to a flurry of research to better understand them.

Table 2-1. Leptons

Particle Name	Symbol	Charge	Mass ($\mathrm{MeV/c^2}$)	Mean lifetime (s)
electron	e	-1	0.511	Stable
muon	μ	-1	105.66	2.197×10^{-6}
tau	τ	-1	1776.99 ± 27	$(2.906\pm10)\times10^{-13}$
electron-neutrino	ν_{e}	0	$< 2 \times 10^{-6}$	Stable
muon-neutrino	ν_{μ}	0	<0.19	Stable
tau-neutrino	ν_{τ}	0	<18.2	Stable

Quarks, shown in Table 2-2, also come in three generations with two flavors per generation, making 6 flavors altogether: up, down, strange, charm, bottom, and top. Quarks are acted on by all three forces. Each quark has an associated anti-quark, and each has a further quantum number denoted *color*; red, blue and green for quarks, with

analogous colors for the anti-quarks. Quarks are always observed bound together into composite particles, called *hadrons*, by a process known as *color confinement*. This produces combinations of quarks that have integer charge. Baryons, such as protons and neutrons, comprise three quarks, each with a different color, yielding a colorless combination. Mesons consist of a quark-antiquark pair also with net zero color.

Quarks have never been observed on their own. When produced at the LHC, quarks hadronize together with the carriers of the strong force, gluons, into observable *jets* of hadrons. These jets, detected and measured in the CMS detector, form one part of the signatures of events under investigation in this thesis analysis. Particular attention will be paid to the jets originating from the hadronization of b-quarks.

Table 2-2. Quarks

Particle Name	Symbol	Charge	Mass ($\mathrm{GeV/c^2}$)
down	d	-1/3	0.35
up	u	+2/3	0.35
strange	s	-1/3	0.5
charm	c	+2/3	1.5
bottom	b	-1/3	4.5
top	t	+2/3	174.2

Charged leptons and quarks are known as *Dirac* particles, meaning they have corresponding antimatter particles. Antimatter is identical to ordinary matter but has the opposite charge. It remains a mystery whether neutrinos have unique antimatter partners, making them Dirac particles, or if they serve as their own antimatter partner.

Terrestrial matter consists of fermions in the first generation: electrons, electron neutrinos, and protons and neutrons which are made of up and down quarks. The other particles in the SM can be artificially produced in particle collisions or occur naturally in upper atmosphere collisions or stellar reactions. However, they are more massive than first-generation particles and therefore have shorter lifetimes. Antimatter is not observed because it annihilates with ordinary matter, and there is overwhelmingly more matter than antimatter in the universe.

2.2 Gauge Bosons

The forces of the SM all exhibit *gauge invariance*, meaning the physical predictions are unchanged when their potentials undergo the gauge transformation shown in Eq. 2–1, where f is an arbitrary scalar function.

$$\phi \rightarrow \phi + \frac{\delta \mathrm{f}}{\delta \mathrm{t}}, \quad \mathrm{A} \rightarrow \mathrm{A} - \Delta \mathrm{f} \tag{2–1}$$

A consequence of gauge invariance is that the force carrier(s) must in theory be massless unless they interact with an underlying scalar field, such as Higgs. The carriers of the weak force, the $\mathrm{W}^{\pm}$ and Z^{0} bosons are known from experiment to be massive force carriers, while photons and gluons are massless. For decades, the unexplained masses of the weak bosons, and fermions, fueled the idea that the Higgs field, or something like it, was probably a necessary part of the SM.

Table 2-3. Spin-1 gauge bosons and spin-0 Higgs boson.

Force	Carrier	Symbol	EM Charge	Mass ($\mathrm{GeV/c^2}$)	Strength	Range
Strong	Gluon	g	0	0	1	10^{-15}
Electromagnetic	Photon	γ	0	$<$6x10^{-17}eV	1/137	inf
Weak	Charged Boson	$\mathrm{W}^{\pm}$	$\pm$1	80.403($\pm$29)	10^{-7}	10^{-18}
Weak	Neutral Boson	Z^{0}	0	91.1876($\pm$21)	10^{-7}	10^{-18}
Higgs Field	Higgs Boson	H	0	125	-	-

2.3 Symmetry

There is intrinsic beauty in the SM by the various symmetries it contains. For each symmetry, there is a correspondingly conserved quantity. Charge conservation before and after an interaction, or C, is a consequence of the matter-antimatter symmetry. Another underlying symmetry is *parity*, or P, in which the chirality of particles is conserved over an interaction. There are other quantities, such as lepton flavor, or baryon number, which are not thought to be fundamentally conserved, but are nonetheless observed rules of nature. Baryon number conservation prohibits the decay of baryons to non-baryonic matter. These observed conservations lead to constraints on the theory of Supersymmetry (SUSY), as described in Chapter 3.

There are exceptions, where these conservation rules are not obeyed, and therefore illuminate some interesting features of the SM. For instance, there are inexplicably CP-violating interactions, which are thought to contribute to the imbalance of matter and antimatter in the universe, but this remains an open mystery. In another example, neutrino oscillations were discovered because of an apparent violation in lepton number, when in fact the neutrinos were oscillating to a different flavor in flight to the detector, leading to the discovery of neutrino oscillation.

2.4 Questions Unanswered by the SM

Despite the incredible predictive power of the SM and the fact that all observed particles are described by its mathematical framework, there remain some unresolved questions for the SM.

- **Matter-Antimatter Imbalance:** Assuming equal amounts of matter and antimatter when the universe began, the only way to have an excess of matter today is for particle interactions to violate barion-number conservation. However, there are not enough known CP-violating interactions to account for this.

- **Dark Matter:** Most of the matter in the universe is *dark*, meaning it only interacts weakly. Neutrinos are the only known SM particle that qualifies as dark matter, but they are not massive enough. There must be additional, unknown weakly-interacting particles in nature to account for dark matter in the universe.

- **Dark Energy:** The universe is expanding at an accelerated rate, which can be achieved if there is an additional energy separate from matter or dark matter in the universe. This dark energy is thought to arise from the cosmological constant, or λ, which was devised by Albert Einstein to grant an energy density to the vacuum of space, nominally his “greatest blunder”. Dark energy would have a very small, but uniform density throughout the universe.

- **Gravity:** To date there is no coherent quantum field theory of gravity, but if it exists, the graviton would have spin 2. Gravity is very weak, with a relative strength of 10^{-40} when compared to the strong force. It is unknown why gravity is so weak.

- **Hierarchy Problem:** As a scalar particle, the Higgs boson requires quantum corrections to the calculation of its mass which involves quadratically divergent integrals. The integrals are therefore capped at an artificial upper limit, Λ_{UV}, to get a finite answer, bounding the energy range in which the Higgs theory is valid. The value of Λ_{UV} should be at least as high as the Plank scale, $\mathrm{M_P} \sim 10^{16}$, where gravity becomes important. However, the large Λ^2_{UV} contributions then have to

be canceled by another term in the Lagrangian, called the *bare mass term*, a contrived means of *fine-tuning* the theory.

SUSY is an attractive theory, for the cancellation of large Λ_{UV}^2 terms of the Hierarchy Problem occurs naturally. Thus, detection of SUSY would solve one of the most basic mysteries of the universe. On the other hand, if SUSY is completely ruled out, this mystery remains completely unanswered without invoking very theoretically unappealing fine-tuning.

CHAPTER 3
SUPERSYMMETRY

3.1 The MSSM

The recent discovery of the Higgs Boson answered the question of why particles have mass, but its existence raises other questions. In the calculation of the Higgs mass, there are loop corrections for each particle to which it couples. These correction terms are quadratically divergent, and the bigger the particle masses, the bigger the divergence. This is known as the *Hierarchy Problem*. Notably, the mass correction terms pertaining to fermion interactions have opposite sign to those from boson interactions. This indicates that one can systematically cancel such terms to all orders in perturbation theory by the introduction of a new symmetry between bosons and fermions, a *supersymmetry*. This symmetry takes the form of an operator, which transforms particles to a supersymmetric partner with spin offset by 1/2. Spin-1/2 fermions have spin-0 superpartners, scalar leptons and scalar quarks, i.e. *sleptons* and *squarks*. Spin-1 gauge bosons and the spin-0 Higgs transform to spin-1/2 superpartners, called *gaugi-nos* and *higgsinos*. These particles, shown in Tables 3-1 and 3-2, constitute the Minimal Supersymmetric Standard Model, or MSSM. Superpartners are denoted with a tilde ($\tilde{}$). [1]

Table 3-1. Fermions and their supersymmetric partners of the MSSM.

Names		spin 0	spin 1/2	$\mathrm{SU(3)_C}$, $\mathrm{SU(2)_L}$, $\mathrm{U(1)_Y}$
squarks, quarks	Q	$(\tilde{\mathrm{u}}_\mathrm{L}\ \tilde{\mathrm{d}}_\mathrm{L})$	$(\mathrm{u}_\mathrm{L}\ \mathrm{d}_\mathrm{L})$	$(\mathbf{3},\ \mathbf{2},\ \frac{1}{6})$
($\times$ 3 families)	$\bar{\mathrm{u}}$	$\tilde{\mathrm{u}}^\star_\mathrm{R}$	$\mathrm{u}^\dagger_\mathrm{R}$	$(\bar{\mathbf{3}},\ \mathbf{1},\ -\frac{2}{3})$
	$\bar{\mathrm{d}}$	$\tilde{\mathrm{d}}^\star_\mathrm{R}$	$\mathrm{d}^\dagger_\mathrm{R}$	$(\bar{\mathbf{3}},\ \mathbf{1},\ \frac{1}{3})$
sleptons, leptons	L	$(\tilde{\nu}\ \tilde{\mathrm{e}}_\mathrm{L})$	$(\nu\ \mathrm{e}_\mathrm{L})$	$(\mathbf{1},\ \mathbf{2},\ -\frac{1}{2})$
($\times$ 3 families)	$\bar{\mathrm{e}}$	$\tilde{\mathrm{e}}^\star_\mathrm{R}$	$\mathrm{e}^\dagger_\mathrm{R}$	$(\mathbf{1},\ \mathbf{1},\ 1)$
Higgs, higgsinos	$\mathrm{H_u}$	$(\mathrm{H}^+_\mathrm{u}\ \mathrm{H}^0_\mathrm{u})$	$(\tilde{\mathrm{H}}^+_\mathrm{u}\ \tilde{\mathrm{H}}^0_\mathrm{u})$	$(\mathbf{1},\ \mathbf{2},\ +\frac{1}{2})$
	$\mathrm{H_d}$	$(\mathrm{H}^0_\mathrm{d}\ \mathrm{H}^-_\mathrm{d})$	$(\tilde{\mathrm{H}}^0_\mathrm{d}\ \tilde{\mathrm{H}}^-_\mathrm{d})$	$(\mathbf{1},\ \mathbf{2},\ -\frac{1}{2})$

Table 3-3 shows how the MSSM gauge eigenstates mix to form mass eigenstates, which are experimentally relevant. The quantity $\mathrm{P_R}$ pertains to R-parity, which is

Table 3-2. Bosons and their supersymmetric partners of the MSSM.

Names	spin 1/2	spin 1	$\mathrm{SU(3)_C}$, $\mathrm{SU(2)_L}$, $\mathrm{U(1)_Y}$
gluino, gluon	$\tilde{\mathrm{g}}$	g	$(\mathbf{8}, \mathbf{1}, 0)$
winos, W bosons	$\tilde{\mathrm{W}}^{\pm}\ \tilde{\mathrm{W}}^{0}$	$\mathrm{W}^{\pm}\ \mathrm{W}^{0}$	$(\mathbf{1}, \mathbf{3}, 0)$
bino, B boson	$\tilde{\mathrm{B}}^{0}$	B^{0}	$(\mathbf{1}, \mathbf{1}, 0)$

described in a later section. The h^0 scalar is the Higgs boson at 125 GeV. The other Higgs eigenstates arise because Higgs couples differently to up and down-type quarks, which transform separately in the supersymmetric extension of the SM.

Table 3-3. MSSM gauge eigenstates and mass eigenstates. So far only the h^0 scalar Higgs boson at 125 GeV has been discovered.

Names	Spin	$\mathrm{P_R}$	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$\mathrm{H_u^0}\ \mathrm{H_d^0}\ \mathrm{H_u^+}\ \mathrm{H_d^-}$	$\mathrm{h}^0\ \mathrm{H}^0\ \mathrm{A}^0\ \mathrm{H}^{\pm}$
squarks	0	-1	$\tilde{\mathrm{u}}_\mathrm{L}\ \tilde{\mathrm{u}}_\mathrm{R}\ \tilde{\mathrm{d}}_\mathrm{L}\ \tilde{\mathrm{d}}_\mathrm{R}$ $\tilde{\mathrm{s}}_\mathrm{L}\ \tilde{\mathrm{s}}_\mathrm{R}\ \tilde{\mathrm{c}}_\mathrm{L}\ \tilde{\mathrm{c}}_\mathrm{R}$ $\tilde{\mathrm{t}}_\mathrm{L}\ \tilde{\mathrm{t}}_\mathrm{R}\ \tilde{\mathrm{b}}_\mathrm{L}\ \tilde{\mathrm{b}}_\mathrm{R}$	(same) (same) $\tilde{\mathrm{t}}_1\ \tilde{\mathrm{t}}_2\ \tilde{\mathrm{b}}_1\ \tilde{\mathrm{b}}_2$
sleptons	0	-1	$\tilde{\mathrm{e}}_\mathrm{L}\ \tilde{\mathrm{e}}_\mathrm{R}\ \tilde{\nu}_\mathrm{e}$ $\tilde{\mu}_\mathrm{L}\ \tilde{\mu}_\mathrm{R}\ \tilde{\nu}_\mu$ $\tilde{\tau}_\mathrm{L}\ \tilde{\tau}_\mathrm{R}\ \tilde{\nu}_\tau$	(same) (same) $\tilde{\tau}_1\ \tilde{\tau}_2\ \tilde{\nu}_\tau$
neutralinos	1/2	-1	$\tilde{\mathrm{B}}^0\ \tilde{\mathrm{W}}^0\ \tilde{\mathrm{H}}_\mathrm{u}^0\ \tilde{\mathrm{H}}_\mathrm{d}^0$	$\tilde{\chi}_1^0\ \tilde{\chi}_2^0\ \tilde{\chi}_3^0\ \tilde{\chi}_4^0$
charginos	1/2	-1	$\tilde{\mathrm{W}}^{\pm}\ \tilde{\mathrm{H}}_\mathrm{u}^+\ \tilde{\mathrm{H}}_\mathrm{d}^-$	$\tilde{\chi}_1^{\pm}\ \tilde{\chi}_2^{\pm}$
gluino	1/2	-1	$\tilde{\mathrm{g}}$	(same)

The theory of SUSY, having been invented previously, gained popularity because of its potential to fix the Hierarchy Problem of Higgs. Under the MSSM, there would be many more particles allowed by nature, and many more possible interactions. If this theory is to offer a real description of nature, it is subject to some constraints based on experimental observation:

Broken symmetry: If SUSY particles exist at the same mass as their SM partners, they would have already been observed at lower energy colliders. SUSY particles must have a much higher mass, excluded by observation below the TeV-scale. Since

SUSY particles would certainly have different mass from SM particles, it is said to be a *broken symmetry*. Eq. 3–1 shows how the theory can be modified to accommodate a broken symmetry. However, to preserve its ability to solve the hierarchy problem, the modification cannot be arbitrary.

$$\mathcal{L} = \mathcal{L}_{\mathrm{SUSY}} + \mathcal{L}_{\mathrm{soft}} \tag{3–1}$$

where $\mathcal{L}_{\mathrm{SUSY}}$ contains the terms of the unbroken MSSM, with terms of order Λ_{UV}^2 that perfectly cancel. $\mathcal{L}_{\mathrm{soft}}$ contains soft SUSY-breaking terms of order $\mathrm{m}_{\mathrm{soft}}^2 \log(\Lambda_{\mathrm{UV}}/\mathrm{m}_{\mathrm{soft}})$, where $\mathrm{m}_{\mathrm{soft}}$ represents the mass difference between SM particles and their superpartners. This scheme will only work if the masses of the lightest SUSY particles do not exceed the TeV-scale.

R-parity: One consequence of assuming a fermion-boson symmetry is that new interactions between quarks and leptons are introduced. As a result, proton decay as well as other lepton (L) and baryon (B) number-violating interactions not experimentally observed, become a theoretical possibility. The proton is known from observation to have a lifetime of at least 10^{34} years, meaning it does not decay. Any new theory has to prohibit proton decay to even be considered, which could be achieved through the insertion of a new conservation rule. L or B conservation alone are insufficient candidates because there are heavily suppressed SM processes which are known to violate L or B number. Theorists have contrived a new rule called *R-parity* conservation, whose calculation is shown in Eq. 3–2.

$$\mathrm{P_R} = (-1)^{3(\mathrm{B}-\mathrm{L})+2\mathrm{s}} \tag{3–2}$$

R-parity is conserved if the product of all particles' $\mathrm{P_R}$ before the interaction equals that of the particles after the interaction. R-parity conservation prohibits proton decay but allows independent L or B non-conservation. As shown in Tab. 3-3, all of the SM particles and Higgs have even parity $(\mathrm{P_R} = +1)$, so by definition, all SM

processes automatically conserve R-parity. All the supersymmetric particles have odd parity ($\mathrm{P_R} = -1$). If R-parity is conserved, only SUSY$\rightarrow$SUSY+SM decays are allowed, and the lightest SUSY particle, or LSP, cannot decay to SM particles via LSP$\rightarrow$SM+SM interactions. However, LSP annihilation to SM particles is allowed, i.e. LSP+LSP$\rightarrow$SM+SM.

It should be noted that experimentalists are looking for both R-parity conserving (RPC) and R-parity violating (RPV) signs of SUSY at the LHC, as mechanism in nature to prohibit proton decay may be very different. In R-parity conserving models, the LSP is likely to be the lightest neutralino, $\tilde{\chi}_1^0$ in Tab. 3-3.

If SUSY takes on the form of the MSSM, there are some amazing predictions that naturally follow with the potential to answer these important open questions:

Dark matter candidate: Dark matter makes up roughly 85% of matter in the universe, unaccounted for in the SM. A new theory accounting for dark matter must contain a weakly-interacting massive particle (WIMP). If SUSY exists, then the LSP, serves as a WIMP candidate. By R-parity conservation, the LSP would never decay to ordinary matter, it could only annihilate with another LSP to produce ordinary matter. As the LSP density in the universe decreases, so does the rate of annihilation, resulting in the so-called *relic density*, which remains relatively stable over time. This relic density of LSP matter could account for the dark matter in the universe.

Grand Unification: There is aesthetic appeal to the notion that the interaction strengths of the weak, strong, and electromagnetic forces may converge exactly at some energy scale, called *Grand Unification Theory*, or GUT. Within the current SM, the inverse coupling constants converge somewhat at 10^{14} GeV, shown by dashed lines in Fig. 3.1. When SUSY particles are introduced with TeV-scale masses, the couplings converge more tightly at 2×10^{16} GeV, or $\mathrm{M_U}$, shown by solid lines in Fig. 3.1, where SUSY particle mass is either 500 GeV (blue) or 1500 GeV (red). The energy of unification is little changed by different SUSY particle masses. It is a striking coincidence

that convergence happens around the Plank scale, $\mathrm{M_P} > 10^{16}$ GeV, where new physics is thought to occur.

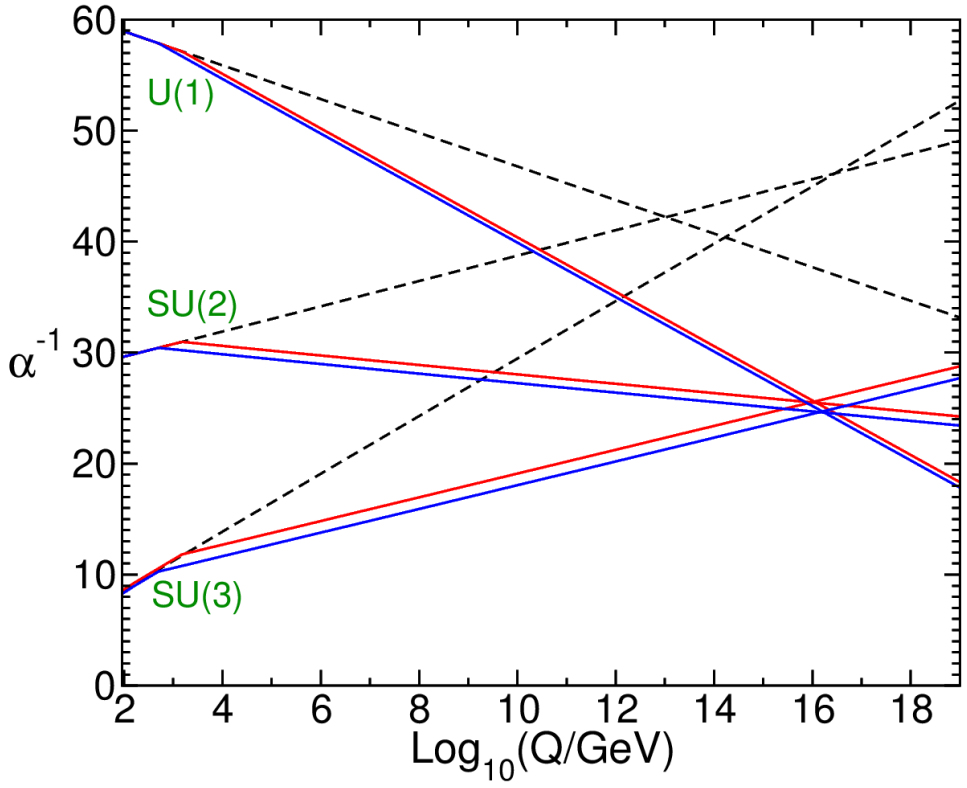


Figure 3-1. Grand unification of the forces.

3.2 SUSY Production at the LHC

The predicted cross-sections of SUSY particle production at 8 TeV as a function of SUSY particle mass are shown in Fig. 3.2 [2]. These cross-sections increase greatly with the center-of-mass energy, and so each step of energy brings a huge increase in the possible SUSY models that could potentially be discovered. Despite the large integrated luminosity of the Fermilab collider, and the mature analyses performed there, turning on the LHC was an exciting moment as new models of SUSY could produce a statistically significant signal even with small datasets. Although the past LHC collision run at 7 TeV showed no hints of SUSY production, even the comparatively minor change of energy to 8 TeV, together with increased luminosity, created new exciting discovery potential.

Strongly interacting sparticles, $\tilde{\mathrm{q}}$ and $\tilde{\mathrm{g}}$, have more degrees of freedom and therefore larger cross-sections as shown in Fig. 3.2. Therefore, they are likely to be discovered before direct electroweak SUSY production, $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_1^0$, called EWKino. EWKino analyses can potentially yield signals with even less hadronic activity in the final state than strongly produced SUSY, and they thus provide the opportunity for complementary searches for SUSY in the LHC environment. EWKino analyses are the subject of a recent University of Florida thesis [3], but this thesis concentrates on searches for SUSY from strongly interacting particles and in particular using the signature of two same-sign leptons.

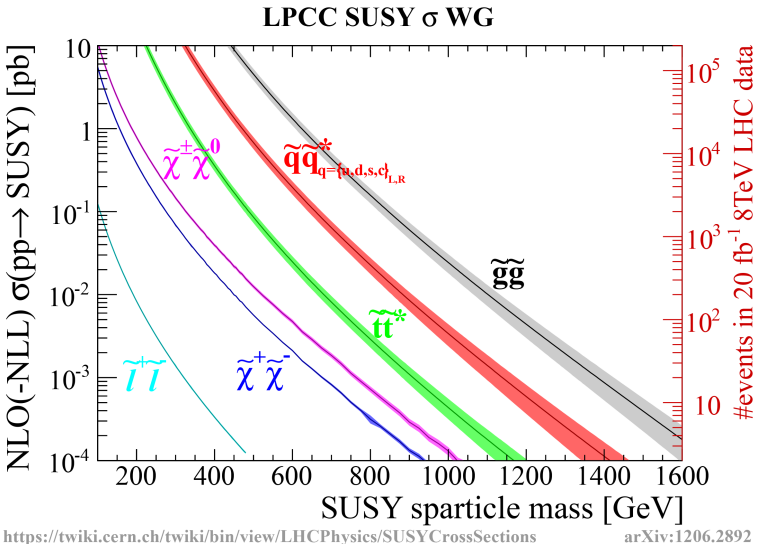


Figure 3-2. SUSY production cross-sections.

3.3 SUSY Production in the Same-sign Dilepton Channel

A typical SUSY production event decaying to same-sign dileptons is shown in Figure 3.3. The SUSY particles, squark and gluino, are produced as a pair as mandated by R-parity conservation. Then they decay according to the mass hierarchy, where strongly-interacting sparticles decay to charginos, which decay to the lightest neutralinos, and leptons. The charged leptons produce an easily-identified signal in the detector, while the neutrinos and LSPs are only seen indirectly as a large imbalance in the final state momenta of the particles, called missing $\mathrm{E_T}$, or $\not\mathrm{E}_\mathrm{T}$.

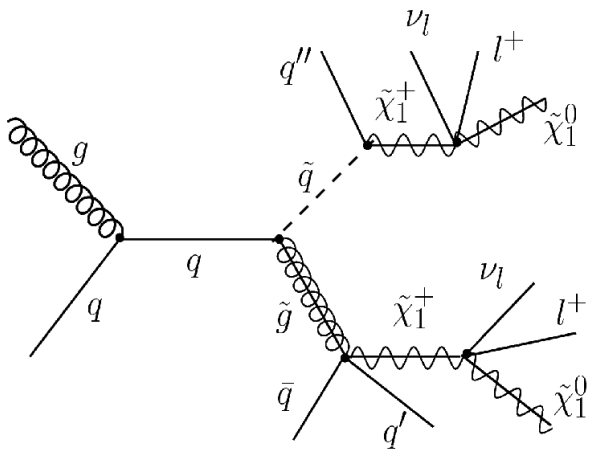


Figure 3-3. SUSY production event decaying to same-sign dileptons.

Such events with two same-sign leptons are rare in SM background processes. They come in the form of $\mathrm{t\bar{t}}$ events, where one lepton decays from a W-boson, called prompt leptons, with a clean, isolated signal in the detector. The second lepton decays from a quark, called nonprompt leptons, occurring within a characteristic hadronic shower, or jet, and thus are not isolated from surrounding detector activity. Same-sign lepton pairs also occur in diboson production events, such as WW and WZ, in which both leptons are prompt, but these SM processes are extremely rare. In many SUSY models, however, same-sign dilepton events are produced in abundance, making this a very sensitive channel to new physics.

3.4 SUSY Models Used in Limit Setting

While the same-sign dilepton signature provides a powerful means to constrain SUSY models, one can further probe attractive models with the added assumption of at least two b-jets. Such SUSY scenarios which feature light stops ($\tilde{\mathrm{t}}$) and sbottoms ($\tilde{\mathrm{b}}$) are theoretically well-motivated [4–6]. SUSY models featuring these scenarios and used in limit-setting in the analysis herein are shown in Figs. 3-4-3.4. Simplified SUSY models used in LHC analyses are further described elsewhere. [7].

In Fig. 3-4, gluinos are directly produced ($\mathrm{pp} \rightarrow \tilde{\mathrm{g}}\tilde{\mathrm{g}}$), and each gluino decays to two top-quarks and a neutralino LSP ($\tilde{\chi}_1^0$) by the following scenarios:

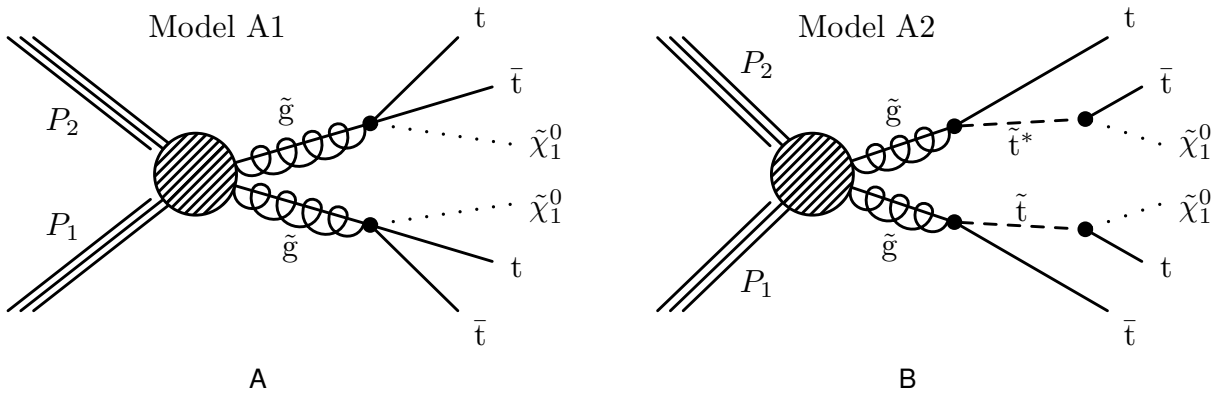


Figure 3-4. Simplified SUSY model diagrams of $\tilde{\mathrm{t}}$ production with virtual $\tilde{\mathrm{t}}$ (A1) and on-shell $\tilde{\mathrm{t}}$ (A2) .

- Model A1, three-body gluino decay mediated by virtual stop: $\tilde{\mathrm{g}} \to \mathrm{t}\bar{\mathrm{t}}\tilde{\chi}_1^0$.
- Model A2, two-body gluino decay to a top-stop pair (on-shell stop): $\tilde{\mathrm{g}} \to \bar{\mathrm{t}}\tilde{\mathrm{t}}_1$ or $\tilde{\mathrm{g}} \to \mathrm{t}\tilde{\mathrm{t}}_1^*$, $\tilde{\mathrm{t}}_1 \to \bar{\mathrm{t}}\tilde{\chi}_1^0$.

These scenarios are abundant when gluinos are lighter than all squarks and stops are the lightest squark. Both scenarios have a signature of $\mathrm{tttt}\tilde{\chi}_1^0\tilde{\chi}_1^0$. The top-quarks decay to W bosons and b-quarks. The W bosons can decay leptonically, giving a final state with at least four leptons, the b-quarks result in at least four b-tagged jets, and $\not\!\!\mathrm{E}_\mathrm{T}$ will result from the LSPs.

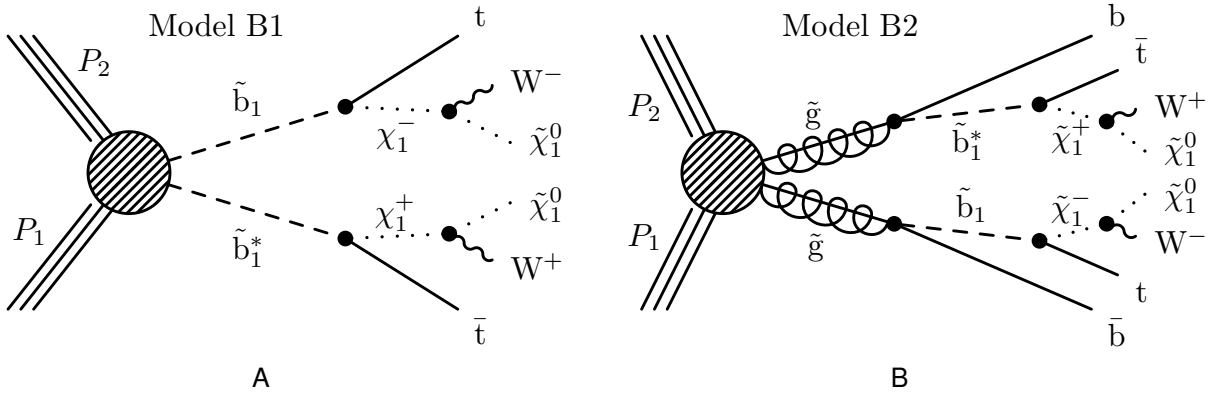


Figure 3-5. Simplified SUSY model diagrams of $\tilde{\mathrm{b}}$ production with direct $\tilde{\mathrm{b}}$ production (B1) and gluino-mediated $\tilde{\mathrm{b}}$ production (B2) .

In Fig. 3-5, two sbottom squarks are produced in the following scenarios:

- Model B1, direct sbottom pair production: pp $\rightarrow \tilde{\mathrm{b}}\tilde{\mathrm{b}}^{*}$
- Model B2, indirect sbottom production from gluino decay: pp $\rightarrow \tilde{\mathrm{g}}\tilde{\mathrm{g}}$, followed by $\tilde{\mathrm{g}} \rightarrow \tilde{\mathrm{b}}\bar{\mathrm{b}}$ or $\tilde{\mathrm{g}} \rightarrow \tilde{\mathrm{b}}^{*}\mathrm{b}$.

These scenarios are abundant if sbottoms are the lightest squark, or lighter than the gluino as in model B1. The sbottom decays to a top-quark and chargino, followed by chargino decay to a W and LSP. Thus the signature of models B1 and B2 both include $\mathrm{ttWW}\tilde{\chi}_1^0\tilde{\chi}_1^0$, with additional b-quarks in model B1. Again, top-quarks decay to W and b-quarks, giving at least four isolated leptons, at least two b-tagged jets, and $\not{\mathrm{E}}_{\mathrm{T}}$ from the LSPs in the final state.

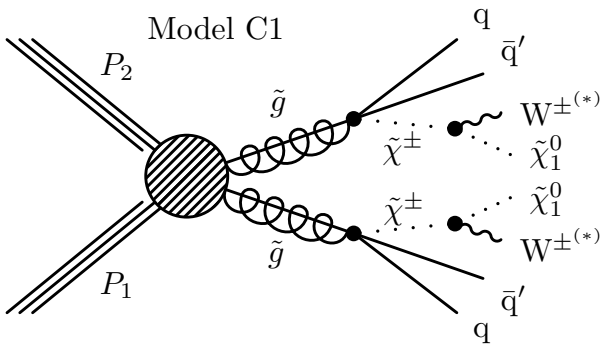


Figure 3-6. Simplified SUSY model diagrams of light squark production (C1).

In Fig. 3.4, three-body gluino decay to same-sign charginos and light quarks is mediated by virtual heavy squarks. Each chargino decays to W and LSP, giving same-sign leptons and $\not{\mathrm{E}}_{\mathrm{T}}$ in the final state.

Parameters such as LSP mass, gluino mass, or intermediate particle mass are varied to generate a model phase space. These models have been simulated with Monte Carlo generated events, varying two parameters at a time to create a $2\mathrm{D}$ grid. A portion of this phase space is ruled out for each model. As a result, limits are set for the masses of SUSY particles within the constraints of these simplified models.

CHAPTER 4
THE LARGE HADRON COLLIDER

4.1 Design

The Large Hadron Collider (LHC) is a proton-proton and heavy ion collider that occupies a circular tunnel 27 km in circumference, 100 m underground, straddling the border between France and Switzerland. There are four beam interaction points, IPs, on the LHC where particles collide. Two are dedicated for general-purpose particle detectors, CMS and ATLAS, used primarily for studdying pp collisions. Another IP is used for studying b-quark physics with the LHCb experiment. The fourth IP is used for studying heavy ion collisions with the ALICE detector. Fig. 4-1 shows the CERN accelerator complex. A complicated chain of particle accelerators boost the protons to their final LHC energy and provide beams to several smaller experiments as well.

The LHC is designed to collide protons at $\sqrt{\mathrm{s}}$ =14 TeV, but after an accident during early running the energy was reduced to a conservative 7 TeV, and was modestly increased to 8 TeV in 2012. The analysis in this thesis is based on the full 2012 dataset at 8 TeV. Since SUSY particles are expected to have masses at the TeV scale, it is important to that the LHC collision energy be as high as possible to make their discovery more likely.

The two proton beams circulate in opposite directions with a separate storage ring for each. Dipole magnets are used for accelerating and steering the beam, while quadrupole magnets are used for focusing the beam. The magnets are made from superconducting cables and require a large cryogenic system to operate at 1.9 K. The LHC has the capacity to circulate over 2,000 bunches of protons in each storage ring, with collisions as frequent as 25 ns apart. Each proton bunch essentially consists of a tightly focused gas of protons. Even though there are $\sim 10^{11}$ protons in each bunch, there are merely tens of hard collisions in each bunch crossing. The design luminosity of pp collisions at the LHC is $\mathrm{L} = 10^{34}\mathrm{cm}^{2}\mathrm{s}^{1}$, given by Eq. 4–1, whose parameters are

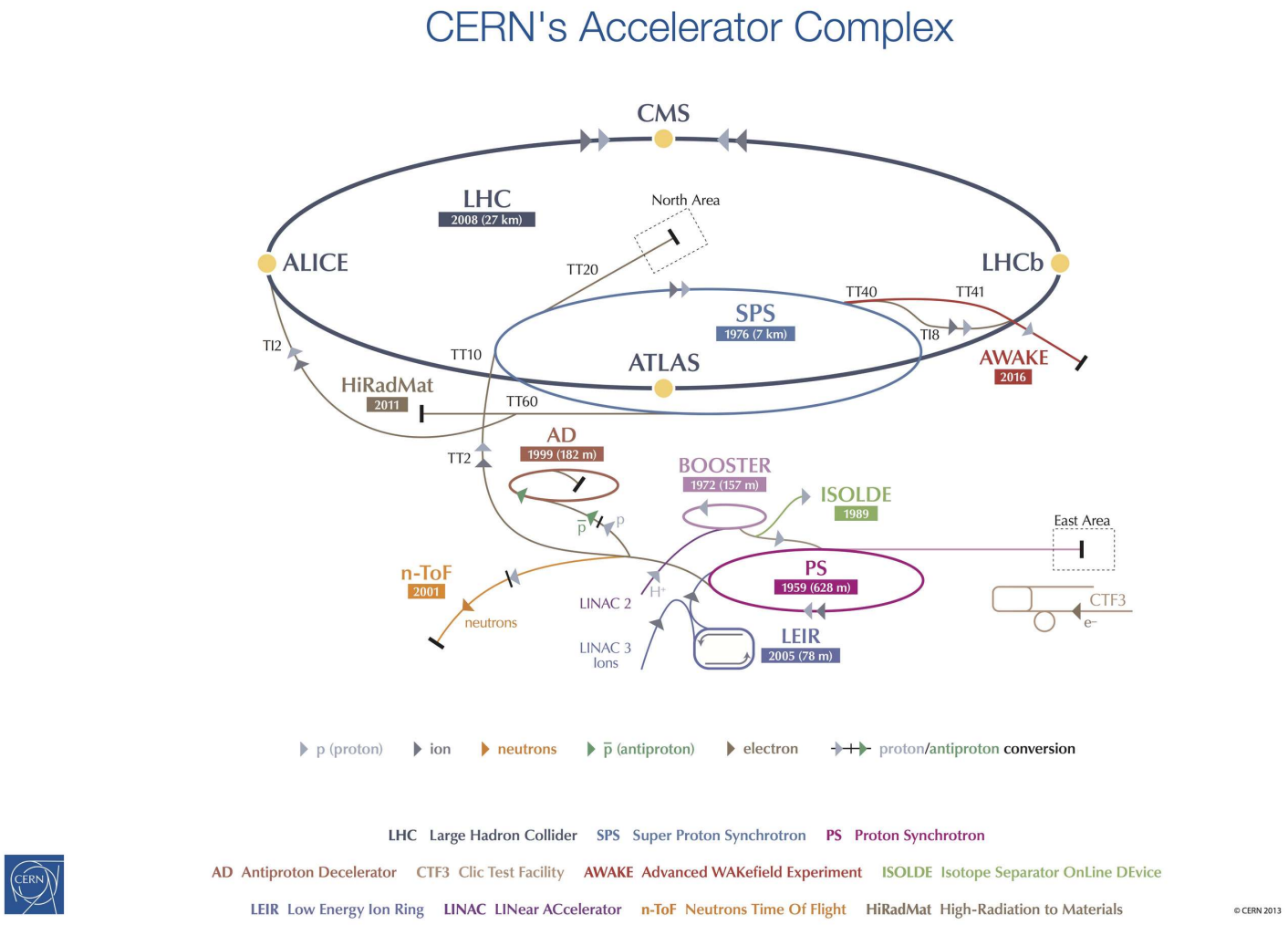


Figure 4-1. CERN accelerator complex.

described in Tab. 4-1. The LHC also collides Pb ions for about two months each year [8].

$$\mathrm{L} = \frac{\mathrm{N_b^2 n_b f_{rev} \gamma_r}}{4\pi \epsilon_n \beta^\star} \mathrm{F} \tag{4–1}$$

Table 4-1. LHC beam parameters.

Parameter	Description	Units	Nominal Value
$\mathrm{N_b}$	protons per bunch		1.15×10^{11}
$\mathrm{n_b}$	bunches per beam		2808
$\mathrm{T_{bs}}$	bunch spacing	ns	25
$\mathrm{f_{rev}}$	orbit frequency	kHz	11.245
γ_r	relativistic boost		7461
ϵ_n	transverse emittance		3.75
$\beta^\star$	beta function at IP	m	0.55

One could instead collide protons and antiprotons, as is done at the Tevatron at Fermilab in Batavia, IL, and only one set of magnetic fields would be needed to store both beams, which is simpler to engineer. However, the difficulty of producing antiprotons along with their shorter beam lifetime results in a large reduction in luminosity. The challenge of building two separate beam lines, but ample luminosity in pp collisions was thought to be a better use of resources at the LHC. The novel design of the LHC magnets has both beam lines going through them, which allows for a small crossing angle and therefore higher luminosity.

4.2 Experiments

CMS - The Compact Muon Solenoid is a high-luminosity, general-purpose detector. The detector design is centered around a 3.8 Tesla superconducting magnet, which is used in the detection of charged particles. CMS also features a high-precision tracker for plotting particle trajectory as well as rapid-response electromagnetic calorimetry for measuring particle energy. In addition, CMS is specifically designed with a highly efficient muon detection system. The physics analyses at CMS include precision SM measurements as well as sensitive searches for new physics such as Higgs and SUSY. More details on the CMS detector can be found in Chapter 5 and elsewhere [9].

ATLAS - A Toroidal LHC ApparatuS (ATLAS) is another general-purpose detector with similar physics analysis priorities to those of CMS, making them competitors. However, ATLAS has a very different design regarding the magnetic field and materials used to build the various sub-detectors, the details of which can be found elsewhere [10]. The benefit of having two completely separate experiments of similar scale with the same physics goals is that there can be independent verification of new physics results, the gold standard of scientific discovery.

ALICE - A Large Ion Collider Experiment (ALICE) is used to study heavy-ion collisions at the LHC. These collisions produce an exotic state of matter, the quark-gluon plasma, which is thought to have existed shortly after the Big Bang occurred. The luminosity delivered to ALICE is several orders of magnitude less than that of the high luminosity experiments, at $\mathrm{L} = 10^{27}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. More information about the design of the ALICE experiment can be found elsewhere [11].

LHCb - The Large Hadron Collider beauty (LHCb) experiment is used to study b-quark physics with the aim to understand why there is more matter than antimatter in the universe. B-quarks are produced in relatively low-energy pp collisions with b-quark momentum primarily along the beam line. Thus the LHCb detector has only one forward spectrometer close to the beam line with a small solid angle of coverage. The LHCb experiment is designed for a smaller luminosity than CMS and ATLAS, at $\mathrm{L} = 10^{32}\mathrm{cm}^{-2}\mathrm{s}^{-1}$. More information about the LHCb experiment can be found elsewhere [12].

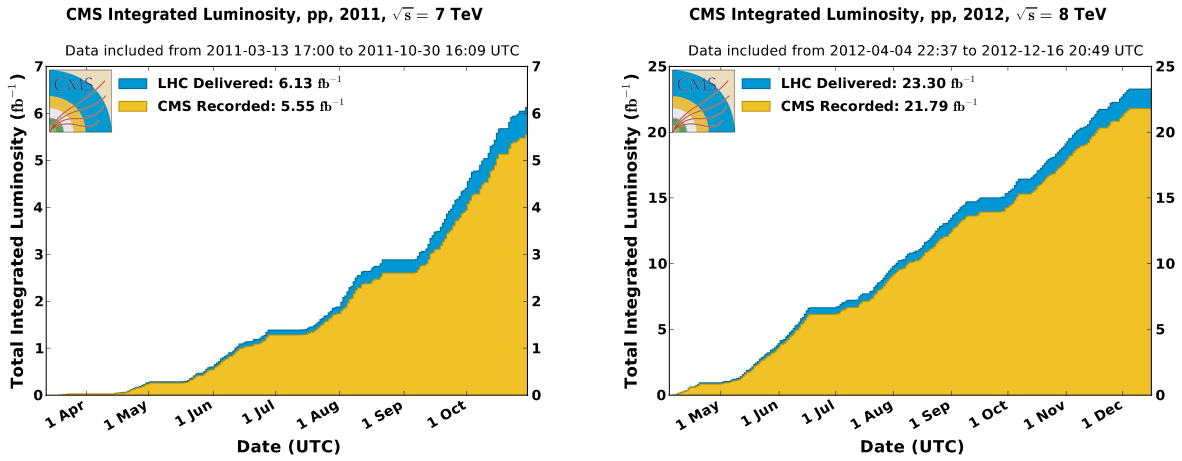
LHCf - The Large Hadron Collider forward (LHCf) experiment consists of two small detectors outside the ATLAS interaction point at $\pm$140 m. The goal of LHCf is to study neutral particles coming from pp collisions to better understand the physics of energetic cosmic rays. More information about the LHCf experiment can be found elsewhere [13].

TOTEM - The TOTal Elastic and diffraction cross-section Measurement (TOTEM) experiment has detectors about the CMS interaction point. TOTEM is used to study

the forward debris of proton collisions to measure the size of the proton as well as the luminosity delivered by the LHC. More information about the design and physics goals of TOTEM can be found elsewhere [14].

4.3 Performance

The integrated luminosity for LHC running in 2011 and 2012 is shown in Fig. 4.3. When part of the CMS detector is found to be malfunctioning during a collision run, the dataset is discarded, thus accounting for the discrepancy between total luminosity delivered by the LHC and that recorded by CMS [15]. This thesis is based on the analysis of 2012 data in Fig. 4.3.



A B

Figure 4-2. LHC integrated luminosity is shown for (A) 2011 and (B) 2012.

CHAPTER 5
THE CMS DETECTOR

The Compact Muon Solenoid (CMS) is a general purpose detector located at Interaction Point 5 of the LHC, about 100 meters underground. It is designed to operate in the high-luminosity environment of the LHC with high-performance tracking, calorimetry, and particle identification. The physics goals include precision SM measurements as well as new physics searches, notably the search for the Higgs particle and SUSY.

CMS has a cylindrical geometry about the beam collision point. The positive z-axis points along the beam line toward the Jura mountains. The x-axis points toward the center of the LHC, and the y-axis points upward. The azimuthal coordinate, ϕ, is measured radially starting from the positive x-axis, while the zenith angle, θ, is measured down from the positive z-axis. The pseudorapidity is defined as $\eta = -\ln(\tan(\theta/2))$, and is preferred because $\Delta\eta$ between two particles is a relativistically conserved quantity, and therefore the outgoing particle density is uniform in pseudorapidity. CMS is divided into the barrel, $|\eta| < 1.1$, and the endcaps, $1.1 < |\eta| < 3.0$.

The sub-detectors of CMS, shown in Fig. 5-1, are arranged such that the least destructive detecting elements are closest to the interaction point. The silicon tracker is the innermost system, followed by the electromagnetic calorimeter (ECAL), then the hadronic calorimeter (HCAL). The HCAL has components both within and outside the superconducting solenoid magnet. Finally the outermost sub-detector is the muon system, which has tracking interleaved with the iron return yoke.

The cylindrical design of the CMS detector arises because of the $3.8\ \mathrm{T}$ superconducting solenoid magnet around which the detector is built. The magnetic field points along the positive z-axis within the solenoid, so a charged particle will bend in ϕ. Magnetic field lines outside the magnet are harnessed and concentrated by an iron return yoke

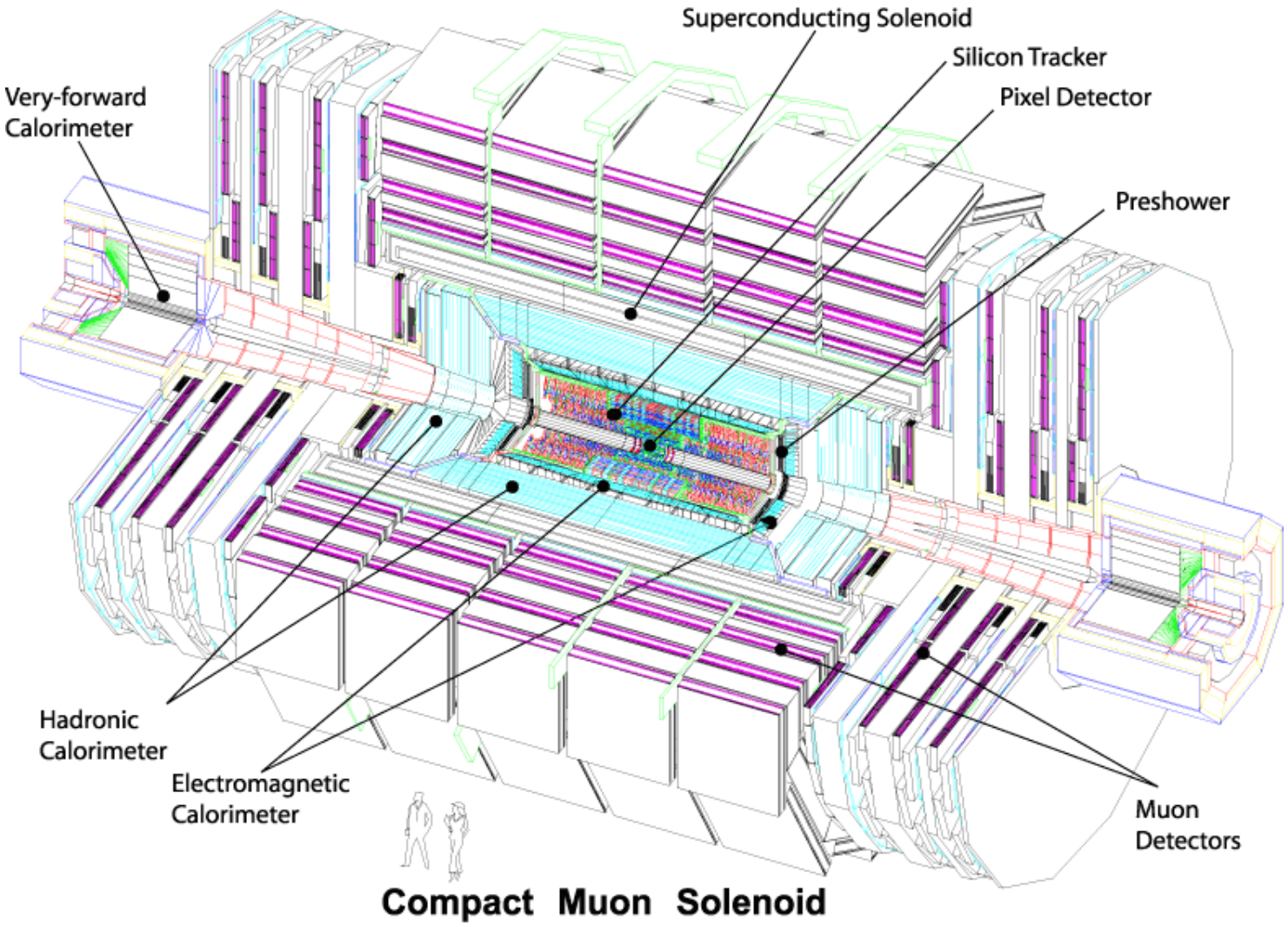


Figure 5-1. Diagram of the CMS detector showing all the sub-detectors to scale.

embedded in the muon system and point along the negative z-axis. The magnetic field allows for the momentum and charge to be measured of charged particles [16].

5.1 Tracking System

Pixel Tracker - The pixel tracker is the innermost tracker to the interaction vertex where particle flux is highest, covering radii from 4.4 cm to 10.2 cm and z from 34.5 cm to 46.5 cm. There are 66 million pixels, at 100 μm $\times$ 150 μm each, with an occupancy of 0.1% per LHC crossing at peak luminosity.

Silicon Microstrips - Silicon microstrip detectors cover the 65 cm $< |\mathrm{z}| <$ 280 cm region. With 9.6 million silicon strips of varying sizes, the strip tracker has an occupancy of about 1-3% at peak luminosity.

As charged particles pass through the tracking system, they interact with the silicon and deposit energy. The resulting ionized charge is read out through electronic channels and form a set of hits which are used to reconstruct the trajectory of the particle. Charged particles follow a helical path, curved in ϕ due to the magnetic field of the solenoid along z. From the curvature of the path, the momentum transverse to the beam line, $\mathrm{p_T}$, can be measured as in Eq. 5–1 , where q is charge, B is magnetic field strength, R is radius of curvature, and θ is the angle to the beam line.

$$\mathrm{p_T} = \frac{0.3\mathrm{qBR}}{\cos\theta} \tag{5–1}$$

Multiple tracks from the same collision are used to reconstruct the primary event vertex. In the case that a short-lived, heavy flavor meson travels some distance from the primary vertex and then decays, a secondary vertex is reconstructed. The transverse impact parameter, $\mathrm{d_0}$, given in Eq. 5–2 , measures the distance of closest approach from the particle track to the primary vertex, where $\vec{\mathrm{V}}$ is the position of the primary vertex. Particles coming from a secondary vertex tend to have a large $\mathrm{d_0}$. For example, $\mathrm{d_0} >$ 0.02 cm was a loose cut used in the same-sign (SS) dilepton analysis to reliably

select particles originating from the primary vertex.

$$\mathrm{d_0} = (\hat{\mathrm{P}}_\mathrm{T} \times \vec{\mathrm{V}}) \cdot \hat{\mathrm{z}} = \frac{\mathrm{V_x} \cdot \mathrm{P_y} - \mathrm{V_y} \cdot \mathrm{P_x}}{|\mathrm{P_T}|} \tag{5–2}$$

Fig. 5-2 shows the combined tracking system resolution of $\mathrm{p_T}$, $\mathrm{d_0}$, and $\mathrm{z_0}$ measurements. Note that $\mathrm{z_0}$ is also referred to $\mathrm{d_z}$ in later chapters. The elements of the tracker are oriented to obtain the most precision in the measurement of x- and y-positions of tracker hits, which is why $\mathrm{d_0}$ has better resolution than $\mathrm{z_0}$.

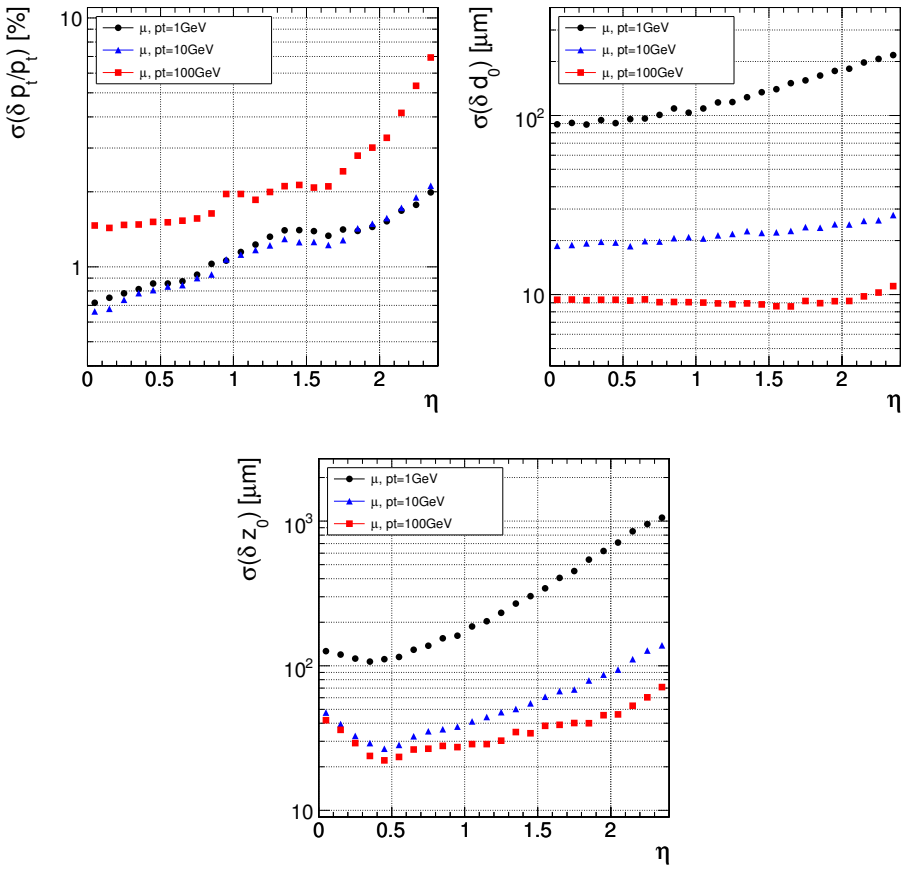


Figure 5-2. Tracking system resolution shown for measurement of $\mathrm{p_T}$ (top left), $\mathrm{d_0}$ (top right), and $\mathrm{z_0}$ (bottom).

5.2 Electromagnetic Calorimeter

The Electromagnetic Calorimeter, or ECAL, is a homogeneous detector made of 68,000 lead-tungstate crystals ($\mathrm{PbWO_4}$). The barrel section has an inner radius of

129 cm and covers the pseudorapidity range $0 < |\eta| < 1.479$. The endcap crystals are 314 cm from the interaction vertex and cover $1.479 < |\eta| < 3.0$. When an electron or photon enters the crystals, it collides with valence electrons, imparting some momentum along the way. The valence electrons will either ionize and proceed to collide with other electrons, or they will be briefly excited to a higher energy level, emitting a photon upon de-excitation. Through successive collisions, an electromagnetic shower is produced in which all the energy of the original electron or photon is absorbed by the crystal and emitted as light. The shower is picked up by photodiodes to measure the total energy of the particle with resolution of 1-3%. One radiation length, $\chi_0 = 0.89\ \mathrm{cm}$, is the distance it takes for an electron to radiate 63% of its energy. In order to capture the entire shower, the ECAL crystals cover about 25.8 χ_0 in the barrel and 24.7 χ_0 in the endcaps. Lead-tungstate crystals were chosen in part because they have a fast response, emitting 80% of the light in 25 ns.

Fig. 5-3 shows the resolution of an energy measurement in an ECAL supercluster (9 crystals) conducted with a test beam of electrons.

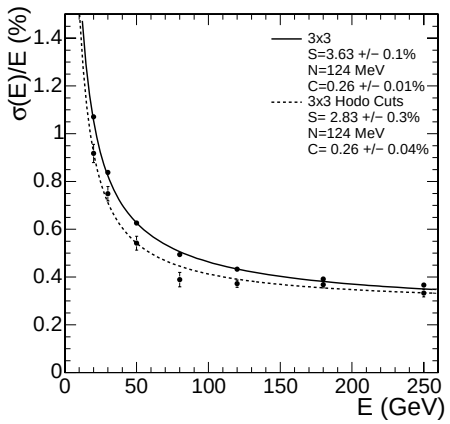


Figure 5-3. ECAL resolution for a test beam shown as a function of electron energy.

5.3 Hadron Calorimeter

The Hadron Calorimeter, or HCAL, is made of steel and brass absorbing material interleaved with scintillator tiles. Hadrons, such as those from jets, enter the HCAL and

interact with the steel and brass layers via nuclear interactions, producing hadronic showers. The hadronic showers produce light in the scintillator tiles which is sampled to measure the energy of the hadrons. The interaction length, λ_{I}, is the mean distance it takes for a hadron to release 63% of its energy via nuclear interactions. The HCAL was not designed to absorb all the hadronic energy, for $\lambda_{\mathrm{I}} = 16\ \mathrm{cm}$. The HCAL interior to the magnet only covers 5.8 λ_{I} within the barrel, and extends to 11.8 λ_{I} outside the magnet. As such, the hadronic energy deposit reported by the HCAL has to be corrected. Furthermore, charged hadrons will shower in the ECAL, as this provides about one interaction length. The inaccuracy of energy measurements in the HCAL dominates the uncertainty in the jet energy measurement, and requires complicated scaling and corrections to be made later in the reconstruction chain. The scaling of jet energy (JES) is a leading source of systematic error in the SS dilepton analysis.

5.4 Muon System

The CMS detector is so-named because it was designed to have excellent muon identification efficiency. Muons are part of the signature of the so-called *golden channel* of Higgs decay, $\mathrm{H} \rightarrow \mathrm{ZZ}^{*} \rightarrow \ell^{+}\ell^{+}\ell^{-}\ell^{-}$, where all the leptons are muons. Muons and electrons are in the SS dilepton event signature, though muons get preferential treatment owing to their cleaner signal in the detector.

Muons travel through most of the detector virtually unencumbered because they are Minimum Ionizing Particles (MIP), meaning muons are not subject to the radiative losses that other particles experience through detector interactions. The MIP nature of muons is due to their larger mass, compared to electrons, resulting in more inertia so they are perturbed less by interactions with the material in the detector. Thus, the muon system is the outermost part of the CMS detector, which allows for a second momentum measurement in addition to the momentum measured by the inner tracker. There are three tracking detectors in the muon system: Drift Tubes (DT) in the barrel, Cathode

Strip Chambers (CSC) in the endcaps, and Resistive Plate Chambers (RPC) in the barrel and endcaps.

Drift Tubes - The drift tubes are laid out in four concentric cylindrical chambers about the magnet in the barrel region. They are layered with the iron yoke, which guides and concentrates the magnetic field outside the solenoid. Each drift cell contains an anode wire in the center of the tube and cathode strips lining the interior. The resulting electric field is used to perform a spacial measurement. When a muon enters the gaseous volume of the drift cell, the gas is ionized by the muon, and the freed electrons drift to the positively charged anion wire. The drift velocity is a known feature of the drift cell, and the drift time is measured, so the location of the incident muon can be reconstructed. These measurements are sensitive to the magnetic field as it can alter the trajectory of the electrons. However, the magnetic field is relatively weak in this part of the detector.

Three of the drift chambers have tubes aligned with the beam line as well as orthogonal to the beam line, allowing for spacial measurements in ϕ and z, respectively. The fourth drift chamber only has tubes to make a ϕ measurement. The drift tubes have a combined spacial resolution of order 100 μm. However, the drift time required can be as much as 400 ns. Note that the CMS detector has to accommodate LHC bunch crossing as frequent as 25 ns.

Cathode Strip Chambers - The Cathode Strip Chambers (CSC) are installed in the endcaps in the range $0.9 < |\eta| < 2.4$ where muon flux is higher and the magnetic field is stronger. The CSCs have a detection principle similar to that of multi-wire proportional chambers which uses alternating orthogonal planes of wires. In each chamber there are 6 anode wire planes with wires along the azimuthal dimension to measure the radial coordinate, and there are 7 orthogonal cathode planes with cathode strips laid out radially to measure the azimuthal coordinate.

When a muon enters a chamber, it passes through a gaseous gap and frees electrons through ionization, similar to the process in a drift tube. The electrons accumulate on the anode wires and induce a differential charge distribution on the neighboring cathode strips. The combination of these CSC hits allows a 2-dimensional r-ϕ measurement to be reconstructed. The performance of CSCs in achieving this 2D measurement is not dependent on the drift time, so they are better suited to operate in a strong magnetic field than are drift tubes. In addition, the CSCs are tilted where appropriate to compensate for the deflection of drifting electrons from the magnetic field.

Resistive Plate Chambers - The Resistive Plate Chambers (RPC) are located in the barrel and endcap regions of the detector. RPCs are parallel plate capacitors that also rely on ionized electrons to drift to an anode, a plate in this case. RPCs have a poor spacial resolution of about 10 mm, but a very rapid response time of 3 ns. Thus RPCs are instrumental in assigning muons detected by drift tubes to the correct bunch crossing in reconstruction. They are also crucial in triggering for muons, as that must happen within nanoseconds of the bunch crossing.

5.5 Triggering

The LHC collides protons in the CMS detector at a rate of 40 MHz, of which only 300 events can be stored each second. This is a firm limit on the bandwidth resulting from the hardware used to buffer and write the events to disk. Most collision events exhibit mundane physics of the SM, such as soft QCD interactions, and can be discarded. Interesting events triggered for storage are often those with hard collisions, where new physics is most likely to result. The LHC is designed to collide protons as rapidly as every 25 ns, so the triggering software must have a rapid response. To this end, full event reconstruction is not performed online. Instead a rough trigger-level reconstruction occurs so the storage decision can be made quickly, while the full event information is kept in a buffer. Triggering constitutes the first step in CMS event selection. The triggering scheme has two stages:

Level-1 Triggers - The L1 triggers reduce the 40 MHz of events to 100 kHz within 3.2 μs. These triggers rely on custom hardware and electronics to facilitate quick online processing of the data from various sub-detectors. The L1 trigger decision is made using coarsely reconstructed event information, such as ECAL energy deposits and track segments. The L1 uses quick algorithms to decide if the events exhibits interesting physics and pass them on to the HLT.

High Level Triggers - The HLTs reduce the 100 kHz of L1-accepted events to 300 Hz. The HLTs make an acceptance decision in about 0.02 to 1 seconds by doing a more refined analysis of L1 event information. These triggers were designed and implemented by analysis experts, according to the physics processes each analysis aims to capture. When developing triggers, the observables of interest are *rate* and *efficiency*. To determine the rate, one must study what events are uniquely triggered by the HLT in question. It is important to monitor the trigger efficiency with respect to individual analysis, to ensure the trigger is not discarding events that would otherwise be selected by that analysis offline when a full event reconstruction is available. It is also possible for inefficiency to arise from a software bug.

SS Dilepton Triggers - Specially designed triggers for the SS dilepton analysis were developed and implemented by the UF group for each dilepton channel: ee, $\mathrm{e}\mu$, and $\mu\mu$. These triggers also require jet activity or missing transverse momentum as one expects to find in SUSY events. Since isolated SS dilepton events are so rare in the SM, these triggers contribute a very small rate to the overall 300 Hz limit, despite having the lowest $\mathrm{p_T}$ threshold for non-isolated leptons in the HLT menu. There were triggers developed independently for Higgs analyses which require isolated leptons but have lower thresholds for jet activity and missing transverse momentum. The Higgs triggers have a large overlap in events triggered with the SUSY SS dilepton analysis triggers. Ultimately, it was decided to use the Higgs triggers to select signal events. However, the requirement of isolated leptons means they are not optimal for use in the background

estimation methods of this thesis. For that, the SUSY triggers are used. The efficiency of these triggers is one contribution to the overall efficiency of this analysis as described by Eq. 1–1. While this efficiency cannot be known exactly, great efforts were made to estimate it and identify inefficiencies whenever possible.

5.6 Physics Objects and Detector Signatures

After an event is triggered by the HLT, it is written to tape. Then the full object reconstruction is performed offline and saved to disk. Reconstruction converts individual detector hits into a physically meaningful picture of collision dynamics. The detector signature of each kind of physics object is shown in Fig. 5-4. The physics objects of interest in the SS dilepton event signature are described below.

Primary Vertex - There are usually dozens of collisions for each bunch crossing, called *pile up*. Each collision is fitted with a vertex with the tracks from the inner tracker. The vertex with the largest $\sum \mathrm{p_T}$ of its final state objects and passing some other validation criteria is chosen to be the primary vertex (PV). All other collisions in the event are ignored. Fig. 5.6 shows the amount of pile up observed at CMS in 2012 with an average of 21 collisions per bunch crossing [15]. It is generally a concern that detector hits from pile up can contaminate the primary event, but that was found to be a negligible effect in this analysis.

Electrons - Electrons are identified in the detector by a track in the tracker and an energy deposit in the ECAL. Electrons are reconstructed and validated by standard CMS software. In this thesis, the selection criteria applied to electrons is knows as the *Medium Working Point* because of the medium selection efficiency of about 95% (though as low as 60% for low-$\mathrm{p_T}$ electrons) and the contamination rate of about 1% [17]. Electrons are the least massive flavor of lepton, and thus lose considerable energy from Bremstrahlung radiation in the tracker, despite its intended design as the least destructive sub-detector. As a result, the curvature of the electron track can be

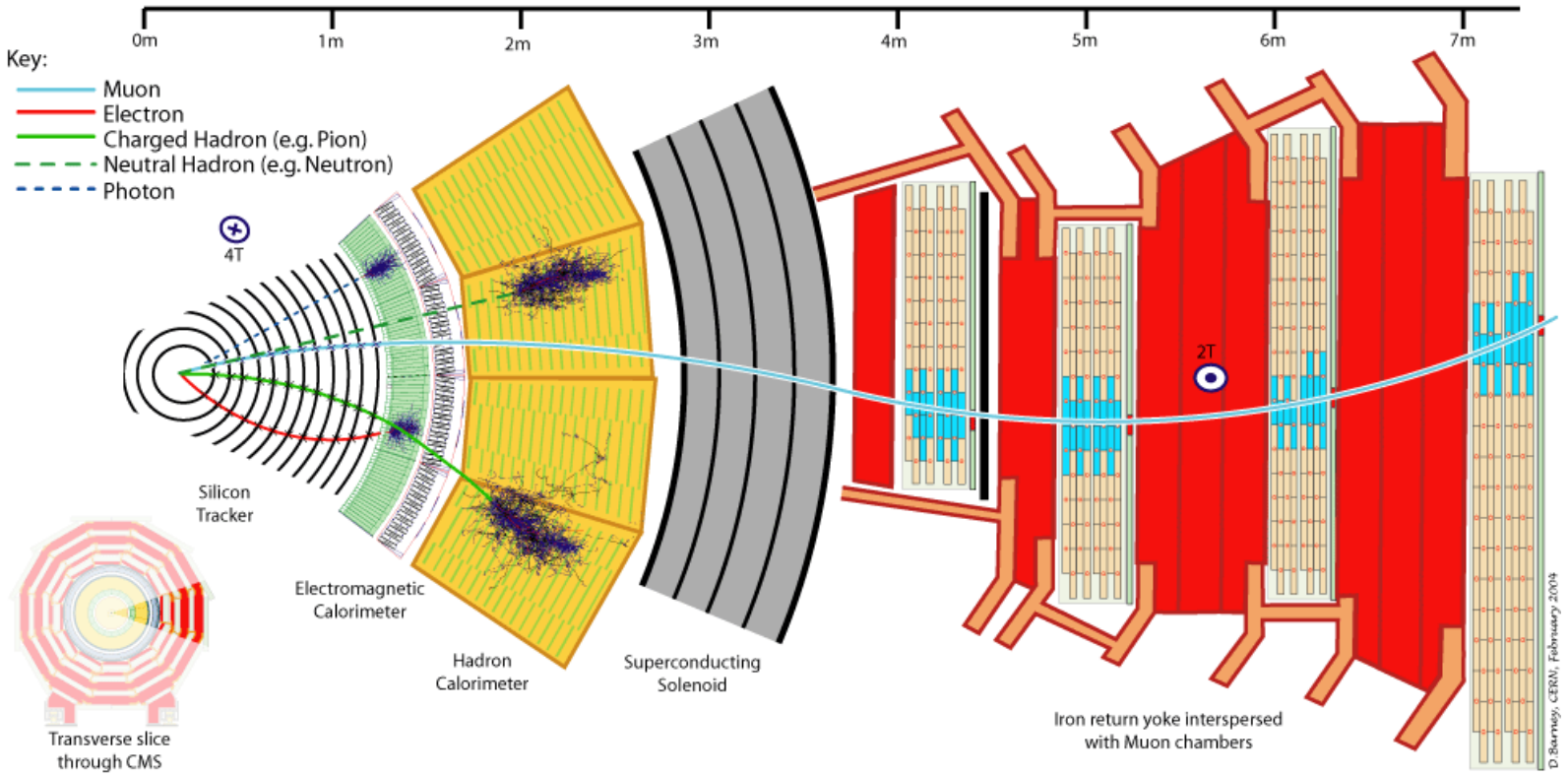


Figure 5-4. Cross-sectional slice of the CMS detector showing the detector signature of different particles.

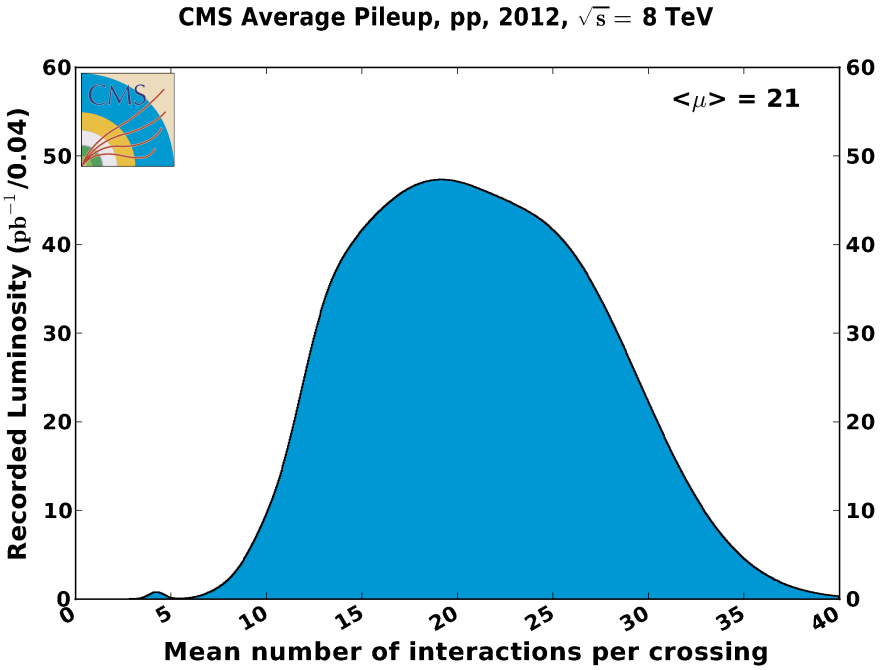


Figure 5-5. LHC pile up is shown for 2012.

altered, subjecting them to charge mis-measurement, which happens at a rate of about $1/10^3$.

Muons - Muons are more massive than electrons, so their signature in the detector differs from electrons. Muons are a minimum ionizing particle (MIP), so they are not subjected to Bremstrahlung losses in the tracker. Muons escape the destructive calorimeter with a relatively small, constant energy deposit of about 2-5 GeV regardless of the actual muon energy. Finally, muons leave a second track in the outer muon detectors with a second momentum measurement. These features combined result in a superbly clean and efficient muon reconstruction at CMS. Muons are selected with criteria known as the *Tight Working Point*, which has an efficiency of 95%, even for low-$\mathrm{p_T}$ muons, and a negligible contamination rate [18]. The charge mis-measurement rate for muons is only $1/10^5$.

Jets - Because of color confinement, quark decays always occur within a shower of hadronization. When hadrons decay, individual quarks are *dressed* by sea quarks,

resulting in a shower of charged and neutral hadrons in the detector. Charged hadrons leave a track in the tracker and energy deposits in the ECAL and HCAL. Neutral hadrons leave only a deposit in the HCAL. The entire hadronic shower typically occurs within a narrow cone. This detector object in its entirety is referred to as a jet. Hadrons can also decay to leptons, so leptons are also contained in a typical jet. The SS dilepton analysis uses a standard CMS jet, found by an anti-$\mathrm{k_T}$ jet-clustering algorithm, with a cone size of $<$0.3. In this analysis, Particle-Flow jets are used, whose reconstruction ideally uses all the available detector information. Jet energy is known to be mis-measured and is corrected in the reconstruction. The disambiguation of signal-like leptons arising from the primary vertex, called *prompt*, and leptons arising from jets, *nonprompt*, is the greatest challenge of this particular SUSY analysis.

Photons - Photons are another well-measured object reconstructed in the CMS detector, based on an energy deposit in the ECAL which has no corresponding track in the tracker. Photons are not part of the event signature of this thesis.

$\not\mathrm{E}_\mathrm{T}$ - The momenta of the final state particles, when summed in the x- and y-directions, should be zero since the colliding partons only have momentum in the z-direction. However, weakly-interacting particles, such as neutrinos or the presumed LSPs, escape detection. Thus, some $\not\mathrm{E}_\mathrm{T}$ is to be expected, but 100 GeV or more would be part of the expected signature of LSPs. The Particle-Flow $\not\mathrm{E}_\mathrm{T}$ algorithm is used in the SS dilepton analysis, which incorporates all the detector information possible.

Clearly, there are many handles in the CMS detector by which SS dilepton events can be selected in the context of some SUSY-based assumptions. In the chapters that follow, the analysis of SS dilepton events is laid out in detail.

CHAPTER 6
VERTEXING

6.1 Motivation

The primary handles for distinguishing prompt, signal leptons from nonprompt, background leptons are cuts on the impact parameter of the lepton tracks, $\mathrm{d_0}$, and the isolation variable, RelIso . Both of these variables are designed to discriminate between prompt leptons, i.e. those coming directly from the primary vertex, and those that come from the decay of, for instance, b and c quarks. The latter background leptons tend to have significant impact parameters (because of the finite path length of the c and b mesons and baryons), and/or be observed in the proximity of other hadrons from the c or b jets.

The RelIso variable and the value of the requirements placed on it have been well-studied by the CMS collaboration and updated several times as appropriate in the last few years. The impact parameter requirement, on the other hand, has been the subject of fewer studies and thus fewer changes. The definition of $\mathrm{d_0}$ in first studies of this process was a measurement with respect to the beam position. This was later modified to be the impact parameter with respect to the primary vertex (PV), as this leads to a more precise measurement. In principle, this could have led to a tighter d0 requirement, but in fact the value of the variable was not changed at that time. In the most recent iteration of the SS dilepton analysis the $\mathrm{d_0}$ cut with respect to the PV was tightened from $|\mathrm{d_0}| < 0.02$ (0.02) cm to $|\mathrm{d_0}| < 0.01$ (0.005) cm for electrons (muons), with efficiencies for prompt and nonprompt leptons shown in Table 6-1.

Table 6-1. Tight $\mathrm{d_0}$ efficiencies with respect to loose $\mathrm{d_0}$ cuts at 0.02 cm are shown for prompt and nonprompt electrons and muons.

Lepton	Prompt	Nonprompt
e	98.10$\pm$0.03%	75.9$\pm$0.1%
μ	98.96$\pm$0.02%	75.5$\pm$0.1%

Clearly tightening the cut significantly reduced nonprompt backgrounds at a rather small cost in efficiency. However, the optimization study for this change was not exhaustive. Here, we investigate whether using more event information as part of the impact parameter requirement can further improve sensitivity.

If a variation of the $\mathrm{d_0}$ cut can be devised which has the same prompt lepton efficiency already deemed acceptable but better nonprompt rejection than the tight $\mathrm{d_0}$ cut, then it would provide greater sensitivity to future versions of the SS dilepton analysis and could be easily approved to replace the tight $\mathrm{d_0}$ cut. The $\mathrm{d_0}$ observable uses the x and y position information of the lepton track. There is perhaps an opportunity to improve the discriminating power of this observable by incorporating z information in the form of a $\mathrm{3D\,d_0}$ cut or by using the x, y, and z track error matrix to form a χ^2 of the compatibility of the dilepton vertex with the event vertex. There is already a very loose cut applied to leptons on $|\mathrm{d_z}| < 0.1$ cm to ensure the track comes from the PV, but this would appear to under-utilize the available information in the z-direction. Future versions of the SS dilepton analysis would have to include the loose $\mathrm{d_0}$ and $\mathrm{d_z}$ cuts, so they are applied throughout the study that follows.

6.2 Selection Criteria

Efficiencies of the cuts being considered are measured with Monte Carlo (MC) events where prompt and nonprompt events are selected by checking the MC truth for leptons coming from W/Z decays.

The leptons selected for this study pass the criteria of the SS dilepton analysis shown in Tables 7-1 and 7-2 and event selection shown in Table 6-2, except for the last row requiring number of leptons. Other event requirements of the SS dilepton analysis, such as requiring a same-sign lepton pair and Z-veto are not applied for the purpose of finding single lepton efficiencies, as it is not necessary to reject these backgrounds in a study of MC generated signal events. For the event efficiencies, the full SS dilepton analysis cuts are applied with the exceptions stated in Table 6-2.

Table 6-2. Event selection for the vertexing study is to apply the full SS dilepton selection with the following exceptions.

Prompt	Nonprompt
T1TTTT	TTJets
Leptons from W/Z	Leptons *not* from W/Z
Loose $\mathrm{d_0}$	Loose $\mathrm{d_0}$
RelIso relaxed	RelIso relaxed
$\mathrm{H_T}$ >250 GeV	$\mathrm{H_T}$ relaxed
$\not\mathrm{E}_\mathrm{T}$ >100 GeV	$\not\mathrm{E}_\mathrm{T}$ relaxed
$\mathrm{N_{jets}} \geq 4$	$\mathrm{N_{jets}}$ relaxed
2 prompt leptons	$\geq$1 nonprompt lepton

The RelIso cut is known to be independent of the cuts being studied, so it has been relaxed to increase statistics in the nonprompt category and thus in prompt events for consistency. Prompt events are required to have hadronic activity with cuts on $\mathrm{H_T}$, $\not\mathrm{E}_\mathrm{T}$, and $\mathrm{N_{jets}}$ to ensure they are similar to the prompt events in our search for gluino/sbottom production.

6.3 New Selection to Consider

For this study, a new object is created by vertexing the same-sign dilepton pair, called the SUSY vertex (SV). The vertex is fit via the Kalman algorithm, a least-squared fit, which uses lepton position and momentum in all three dimensions. This is the same algorithm used for finding the primary vertex of the event. The `C++` code for vertexing two leptons is found in Appendix A.

The quality of a vertex fit, measured by χ^2, is based on the error matrices of its constituent tracks. The χ^2 of the SUSY vertex, denoted here as $\mathrm{SV}\ \chi^2$, is studied as a possible event cut. On average, the χ^2 of a 1D measurement is 1, and for a masurement with 3 degrees of freedom, as with vertex position, average χ^2 is 3. Compatibility is the χ^2 difference between two measured quantities, given by Eq. 6–1 for two 3D measurements, where $\vec{\mathrm{X}}$ is the position and σ is the error matrix. Another event cut studied here is the χ^2 of the compatibility of the SUSY vertex with the primary vertex,

denoted $\mathrm{SV\text{-}PV}\ \chi^2$. The code for compatibility of two vertices is in Appendix B.

$$\chi^2 = (\vec{\mathrm{X}}_1 - \vec{\mathrm{X}}_2)^{\mathrm{T}}(\sigma_1 + \sigma_2)^{-1}(\vec{\mathrm{X}}_1 - \vec{\mathrm{X}}_2) \tag{6–1}$$

Similarly one can find the χ^2 of the compatibility of a single track with a vertex, which is applied in this study as a lepton cut, called $\mathrm{Tk\text{-}PV}\ \chi^2$.

Further, a $\mathrm{Combo\ Cut}$ is defined for the purpose of this study as the application of both the $\mathrm{SV}\ \chi^2$ and $\mathrm{SV\text{-}PV}\ \chi^2$ cuts, optimized to reject as much background as possible, while fixing the prompt lepton efficiency equal to that of the tight $\mathrm{d_0}$ cut. All the cuts are listed below, and each is either applied as a cut on leptons or a cut on events.

Lepton cuts:

- $\mathrm{3D\ d_0}$ - Incorporates the z-position of the lepton track into the existing $\mathrm{d_0}$ observable.
- $\mathrm{Tk\text{-}PV}\ \chi^2$ - Compatibility of the lepton track with the PV.

Event cuts:

- $\mathrm{SV}\ \chi^2$ - Quality of the SV fit, found by vertexing the SS dilepton pair.
- $\mathrm{SV\text{-}PV}\ \chi^2$ - Compatibility of the SV and PV.
- $\mathrm{Combo\ Cut}$ - Simultaneous cuts on $\mathrm{SV}\ \chi^2$ and $\mathrm{SV\text{-}PV}\ \chi^2$.

6.4 Results

The nonprompt single lepton efficiencies of the cuts being studied are shown in Table 6-3. The cuts have been chosen to have the same prompt lepton efficiency as the tight $\mathrm{d_0}$ cut, as this efficiency is considered satisfactory. This allows for a direct comparison of the cuts being considered to those used in the official analysis. The proposed $\mathrm{3D\ d_0}$ cut appears to have a significant advantage in background rejection over those already in use.

Table 6-3. Single lepton efficiencies for vertexing study are shown for nonprompt leptons with respect to loose $\mathrm{d_0}$.

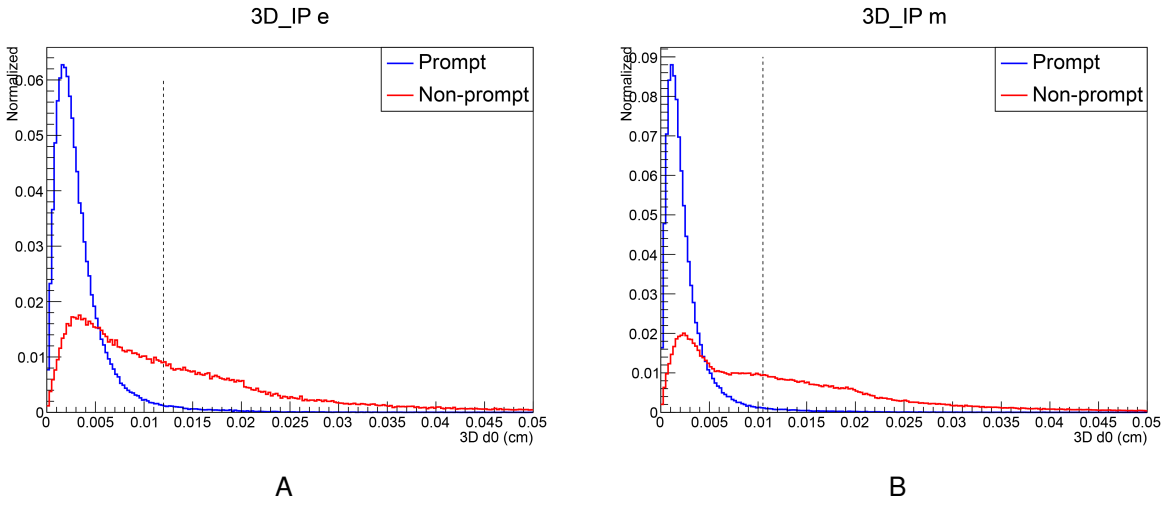
	Tight $\mathrm{d_0}$	$\mathrm{3D\ d_0}$	$\mathrm{Tk\text{-}PV}\ \chi^2$
e	75.9±0.1%	61.9±0.2%	67.5±0.1%
μ	75.5±0.1%	66.9±0.1%	76.9±0.1%

Nonprompt event efficiencies are shown in Table 6-4. These events pass the full SS dilepton analysis event selection, which appears to have reduced the benefit of the $3\mathrm{D}\,\mathrm{d}_0$ cut to some extent. However, it is still the best option of those presented in Table 6-4 for improving the analysis in the future and warrants consideration.

Table 6-4. Same-sign dilepton event efficiencies for vertexing study are shown for nonprompt events with respect to loose d_0.

	Tight d_0	$3\mathrm{D}\,\mathrm{d}_0$	SV χ^2	SV-PV χ^2	Combo Cut
$\mu\mu$, eμ, ee	56.1±0.2%	51.8±0.4%	62.2±0.2%	55.6±0.2%	53.4±0.2%

Fig. 6-1 shows the prompt and nonprompt $3\mathrm{D}\,\mathrm{d}_0$ distributions for electrons and muons. The cuts used were $|3\mathrm{D}\,\mathrm{d}_0| < 0.0120(0.0105)$ cm for electrons(muons), which have the same prompt event efficiency as tight d_0 with respect to loose d_0.



A B

Figure 6-1. $3\mathrm{D}\,\mathrm{d}_0$ distribution plots are shown for (A) electrons and (B) muons, where prompt leptons are shown in blue and nonprompt are in red. $3\mathrm{D}\,\mathrm{d}_0$ cuts on leptons are tuned to match the signal event efficiency of the tight d_0 cut.

Figure 6-2 shows how an optimal Combo Cut was achieved and that it does improve nonprompt event rejection over the tight d_0 cut. In the distribution plots (A and C), the SV χ^2 is plotted vs. SV-PV χ^2, showing that they are largely uncorrelated observables. In the efficiency plots (B and D), each bin gives the efficiency of applying a Combo Cut

there, i.e. the integral of bins both below and to its left in the distribution plot divided by total entries in the distribution plot.

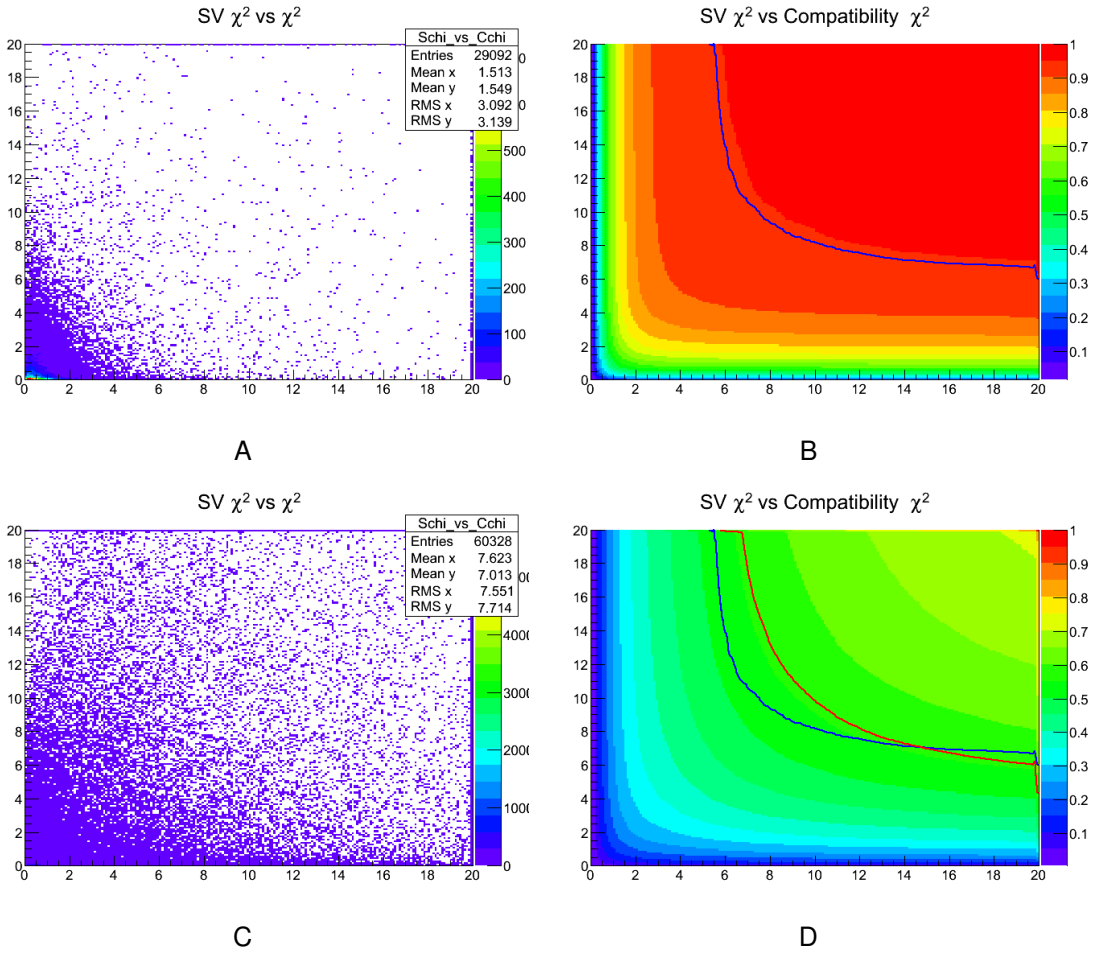


Figure 6-2. $\mathrm{Combo\,Cut}$ distribution and efficiency plots are shown for prompt events (A and B) and nonprompt events (C and D). The blue contour in B is the locus of candidates for the $\mathrm{Combo\,Cut}$ which preserve prompt event efficiency and is copied to D. The red contour in D is the nonprompt efficiency of the tight $\mathrm{d_0}$ cut for comparison.

The blue contour in B shows where one could apply the $\mathrm{Combo\,Cut}$ to achieve the same signal event efficiency of the tight $\mathrm{d_0}$ cut. The red contour in D shows the nonprompt efficiency of the tight $\mathrm{d_0}$ cut. Adopting a $\mathrm{Combo\,Cut}$ on the blue contour would improve nonprompt event rejection, and the greater the distance between the red and blue contours, the bigger the improvement. However, at most, the nonprompt event efficiency can be reduced from 56.1% to 53.4% as shown in Table 6-4, which is

still not as good as the $3\mathrm{D}\,\mathrm{d}_0$ cut at 51.8%. Note that the efficiency plots include the overflow events in the top row and last column, showing the result of applying each cut individually, for which numbers are readily available in Table 6-4.

In conclusion, we have studied whether large improvements can be made to the efficacy of the impact parameter cuts by including more information. We find that there is scope for some improvement, though it is by no means dramatic. In particular, moving from the currently-used $2\mathrm{D}\,\mathrm{d}_0$ cut to the $3\mathrm{D}\,\mathrm{d}_0$ cut can reduce the background significantly while preserving the same signal efficiency. We do not find evidence that more complicated requirements lead to further improvements.

It should be noted that incorporating any such new cut into the entire analysis would take many months of work and would have delayed the publication. This is due largely to the complicated manner in which the backgrounds are calculated in the full analysis chain. However, future analyses could be built with such a cut from the beginning and thus be better-optimized to reject nonprompt backgrounds.

CHAPTER 7
SAME-SIGN DILEPTON SEARCH FOR NEW PHYSICS

7.1 Motivation

The analysis presented in this chapter was one of three major efforts at CMS to search for new physics in same-sign (SS) dilepton events requiring $\not\mathrm{E}_\mathrm{T}$ and b-jets. All three efforts were combined into a common result and published by the CMS collaboration using the full 2012 dataset with 19.5 fb^{-1} of integrated luminosity [19]. An earlier version of this combined result was also published using the first 10.5 fb^{-1} of 2012 data [20].

The analysis of this thesis is performed with two concurrent search strategies, based on independent sets of triggers: one with historically soft leptons (e, μ) corresponding to a $\mathrm{p_T}$ threshold of 10 GeV, denoted low-$\mathrm{p_T}$, and one with hard leptons corresponding to a $\mathrm{p_T}$ threshold of 20 GeV, denoted high-$\mathrm{p_T}$. The search was designed to cover a variety of SUSY models and to be sensitive to final states with different kinematics by using search regions in bins of $\mathrm{N_{jets}}$, $\mathrm{N_{b-jets}}$, $\mathrm{H_T}$, and $\not\mathrm{E}_\mathrm{T}$.

Events with two SS leptons provide a very clean signature in which to search for new physics as the SM background is very low. SS dilepton events where the leptons are prompt, such as leptons from vector boson decay, are very rare in SM processes. A search strategy was designed based upon a series of assumptions on the production of SS dileptons which are made in the context of, but not limited to, SUSY. The assumptions have the following experimental consequences:

- First assume the lightest SUSY particle (LSP), such as the lightest neutralino, $\tilde{\chi}_1^0$, is stable in the case of R-parity conservation and weakly interacting. Assign mass scale $\mathrm{m_C}$ to the LSP. The LSP would pass through the detector without leaving any signal. Its signature would be observed indirectly due to an imbalance of the measured $\mathrm{p_T}$ of outgoing particles, $\not\mathrm{E}_\mathrm{T}$. Thus searching for events with $\not\mathrm{E}_\mathrm{T}$ is an important part of the search strategy.
- Assume SUSY is strongly produced via squarks and gluinos with mass scale $\mathrm{m_A}$. The large cross-section of these processes compared to electroweak SUSY processes (EWKino) makes their discovery in early data more likely. The decay

of squarks and gluinos to colorless SUSY particles leads to jets in the final state. If the top-like squarks are the lightest generation, the experimental signature can be narrowed further by specifically looking for b-jets in the final state. Because of the large mass of the b-quark, the production rates for b-jets is lower than for the lighter flavors. Thus, looking for SUSY events that include b-jets are intrinsically less likely to have major backgrounds.

- Finally, assume a chargino that couples to quarks and squarks with mass scale $\mathrm{m_B}$, such that $\mathrm{m_A} > \mathrm{m_B} > \mathrm{m_C}$. This mass hierarchy makes possible the following decay of SUSY particles: $\tilde{\mathrm{q}} \to \tilde{\chi}^{\pm}\mathrm{q} \to \mathrm{l}^{\pm}\nu\tilde{\chi}_1^0\mathrm{q}$. SUSY particles are pair produced, so two such decay chains would lead to a pair of same-sign leptons in the final state, at least two jets, and $\not\mathrm{E}_\mathrm{T}$ on the scale of $2\mathrm{m_c}$.

The total amount of energy available to leptons depends on the mass splitting, $\Delta\mathrm{m_{BC}}$, between charginos and neutralinos. The energy available to jets depends on the mass splitting, $\Delta\mathrm{m_{AB}}$, between squarks/gluinos and charginos, which shows up in the event selection as a $\mathrm{p_T}$ threshold on jets as well as the sum of jet $\mathrm{p_T}$. This last quantity is given the label $\mathrm{H_T}$. The low-$\mathrm{p_T}$ search has a low lepton $\mathrm{p_T}$ threshold but large minimum $\mathrm{H_T}$ and $\not\mathrm{E}_\mathrm{T}$, while the reverse is true for the high-$\mathrm{p_T}$ search.

The two leading sources of backgrounds to our signal searches are rare SM processes (e.g. $\mathrm{t\bar{t}V}$ and diboson production), and $\mathrm{t\bar{t}}$ events. The former can give two same-sign, prompt leptons, but fortunately have low cross-sections. Events from $\mathrm{t\bar{t}}$ production can mimic the signal in one of two ways: (1) one of the two leptons is actually nonprompt and comes from a hadronic jet (most likely a b-jet), while the other lepton is prompt and comes from a W-decay and (2) both leptons come from opposite-sign (OS) W decays, but the charge is mis-measured due to destructive interactions with the tracking detector. A third source of background is from charge mis-measurement of electrons, though this background is very small in relation to $\mathrm{t\bar{t}}$ and rare SM processes. A fourth, and also small, background source comes from QCD production in which both leptons are nonprompt.

One approach to estimating the number of SM predicted background events is to simulate these processes with the help of Monte Carlo (MC) event generators. MC generated events take into account the physics of particle interactions to generate

realistic collision events and their decay products. Then the the outgoing particle kinematic information is sent through a simulation of the CMS detector and the manner in which it identifies particles. Events with diboson or $\mathrm{t\bar{t}}$ plus associated vector boson production must be estimated with the help of MC simulations. These processes constitute an *irreducible* background as both leptons are prompt, so there are no good handles to distinguish them from signal events. Fortunately, such events have small cross-sections and are thus expected to be rare.

There are inevitably differences between MC generated events and actual collisions, so this method of estimating backgrounds is to be avoided if one can instead make use of the characteristics of the real data in control regions, chosen to be similar to, but distinct from, the regions of phase-space under investigation. This method is known in CMS as a data-driven background estimate. The success of a data-driven method depends on establishing adequate control regions to serve as sideband and to measure the rate at which a particular signature is faked in the detector. It is also common to check the performance of the data-driven method against MC predictions, called a *closure test*. For this thesis, two data-driven methods were used. One method is for estimating the backgrounds from nonprompt leptons, the BTag&Probe method, whether they arise from one or two nonprompt leptons, as in $\mathrm{t\bar{t}}$ or QCD processes, respectively. A second data-driven method is used to estimate backgrounds from charge mis-measurement.

7.2 Physics Objects and Discriminating Variables

The main physics objects employed by this analysis are muons, electrons, jets, and missing transverse momentum ($\not\mathrm{E}_\mathrm{T}$). In this section detailed descriptions are given of the reconstruction methods used to identify and characterize these objects. All the techniques employed are in accordance with the ratified standard objects defined and agreed upon by the whole CMS collaboration.

7.2.1 Leptons

The primary handles for distinguishing prompt, signal-like leptons from nonprompt, background-like leptons are RelIso , a parameter that characterizes the extent that the lepton is isolated from other particles, and impact parameter, i.e. $|\mathrm{d}_{0,\mathrm{pv}}|$. In general, prompt leptons originating from W or Z decay produce an isolated signature in the detector, meaning there is little detector activity in the immediate vicinity of the lepton. On the other hand, leptons from background sources, such as the decay products of hadrons (nonprompt leptons), are part of hadronic jets, and so are typically accompanied by a considerable number of charged and neutral products.

The degree of isolation is measured *relative* to the lepton $\mathrm{p_T}$ by the RelIso variable, which identifies detector activity (tracks and calorimeter deposits) in a cone about the lepton as a fraction of the lepton $\mathrm{p_T}$. The Particle Flow calculation of RelIso is given by Eqs. 7–1 and 7–2

$$\mathrm{RelIso}(\mu) = \frac{\sum_{\mathrm{i}} \mathrm{p}_{\mathrm{T,i}}^{\text{Ch Hadron}} + \mathrm{Max}\left(0, \sum_{\mathrm{i}} \mathrm{E}_{\mathrm{T,i}}^{\text{N Hadron}} + \sum_{\mathrm{i}} \mathrm{E}_{\mathrm{T,i}}^{\text{Photon}} - 0.5 \times \delta\beta\right)}{\mathrm{p}_{\mathrm{T}}^{\ell}} \tag{7–1}$$

$$\mathrm{RelIso(e)} = \frac{\sum_{\mathrm{i}} \mathrm{p}_{\mathrm{T,i}}^{\text{Ch Hadron}} + \mathrm{Max}\left(0, \sum_{\mathrm{i}} \mathrm{E}_{\mathrm{T,i}}^{\text{N Hadron}} + \sum_{\mathrm{i}} \mathrm{E}_{\mathrm{T,i}}^{\text{Photon}} - \rho \times \mathrm{area}_{\mathrm{eff}}\right)}{\mathrm{p}_{\mathrm{T}}^{\ell}} \tag{7–2}$$

where β and ρ are correction factors and the sums are performed within a cone of aperture $\Delta\mathrm{R} = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ around the lepton candidate direction. The RelIso variables have been exhaustively studied to be optimum choices for differentiating between leptons from prompt and nonprompt sources.

In addition, prompt leptons are reconstructed very near the primary vertex and thus have a small $|\mathrm{d}_{0,\mathrm{pv}}|$, distance of closest approach, in comparison to nonprompt leptons. The $|\mathrm{d}_{0,\mathrm{pv}}|$ observable is described in more detail in Chapter 5.

The reconstruction level selection criteria of signal muons and electrons are provided in Tables 7-1 and 7-2, respectively. The utility of the additional, *loose* lepton selection criteria in Tables 7-3 and 7-4 are explained in Sec. 7.4.

Table 7-1. Muon selection requirements for the high-p_T (low-p_T) search

Observable	Value or Range
Id	Particle Flow && Global
p_T	$> 20(10)$ GeV
$\|\eta\|$	< 2.4
χ^2/ndof	≤ 10
# silicon layers	> 5
# Valid SA Hits	> 0
# matched muon station	> 1
$\|d_{0,pv}\|$	< 0.005
$\|d_{z,pv}\|$	< 0.1
Ecal/Hcal Non-MIP Veto	$\leq 4/6$ GeV
RelIso(μ)	< 0.1

Table 7-2. Electron selection requirements for the high-p_T (low-p_T) search

Observable	Value or Range
Missing pixel hits	0
$\sigma_{i\eta i\eta}$ (B/E)	$< 0.01/0.03$
$\Delta\phi_{In}$ (B/E)	$< 0.06/0.03$
$\Delta\eta_{In}$ (B/E)	$< 0.004/0.007$
H/E (B/E)	$< 0.1/0.075$
Seed	Ecal-Driven
p_T	$> 20(10)$ GeV
$\|\eta\|$	< 2.4, $\notin [1.4442, 1.566]$
$\|d_{0,pv}\|$	< 0.01
$\|d_{z,pv}\|$	< 0.1
RelIso(e)	< 0.09
ΔR(e, μ)	> 0.1
$\|1/E - 1/p\|$	> 0.05
charge consistency among CTF, GSF and SuperCluster	

The selection criteria in Tables 7-1 and 7-2 are identical to those published previously [20] with the exception of a tighter $|d_{0,pv}|$ for leptons to further suppress the background from nonprompt leptons. The effect of tightening $|d_{0,pv}|$ was studied for prompt and nonprompt leptons and is shown in Figure 7-1. Other strategies that could

Table 7-3. Muon loose selection requirements

Observable	Value or Range
Id	Particle Flow && (Global || Tracker)
$\mathrm{p_T}$	> 10 GeV
$\lvert\eta\rvert$	< 2.4
$\mathrm{RelIso}(\mu)$	< 0.2

Table 7-4. Electron loose selection requirements

Observable	Value or Range
$\sigma_{\mathrm{i\eta i\eta}}$ (B/E)	$< 0.01/0.03$
$\Delta\phi_{\mathrm{In}}$ (B/E)	$< 0.8/0.7$
$\Delta\eta_{\mathrm{In}}$ (B/E)	$< 0.007/0.01$
$\mathrm{H/E}$ (B/E)	$< 0.15/-$
$\mathrm{p_T}$	> 10 GeV
$\lvert\eta\rvert$	< 2.4
$\lvert\mathrm{d_{0,pv}}\rvert$	< 0.04
$\lvert\mathrm{d_{z,pv}}\rvert$	< 0.2
$\mathrm{RelIso(e)}$	< 0.2

be used for reducing the background from nonprompt leptons are explored in Chapter 6.

7.2.2 Jets and MET

Jets and the missing transverse energy are reconstructed using Particle Flow objects. The former are corrected for energy scale using the L1 correction. Particle flow Jets are corrected with L1FastL2L3 and L2L3 residuals for data and without residual correction for the simulations. The details of the jet selection are provided in Table 7-5. In order to characterize the total event hadronic activity in a generic way, we appeal to the $\mathrm{H_T}$ observable, which is simply the scalar transverse momentum sum of the selected jets.

Table 7-5. Jet selection requirements

Observable	Value or Range
$\mathrm{p_T}$	> 40 GeV
$\lvert\eta\rvert$	< 2.4
Id	Loose
ΔR(jet,ℓ)	> 0.4

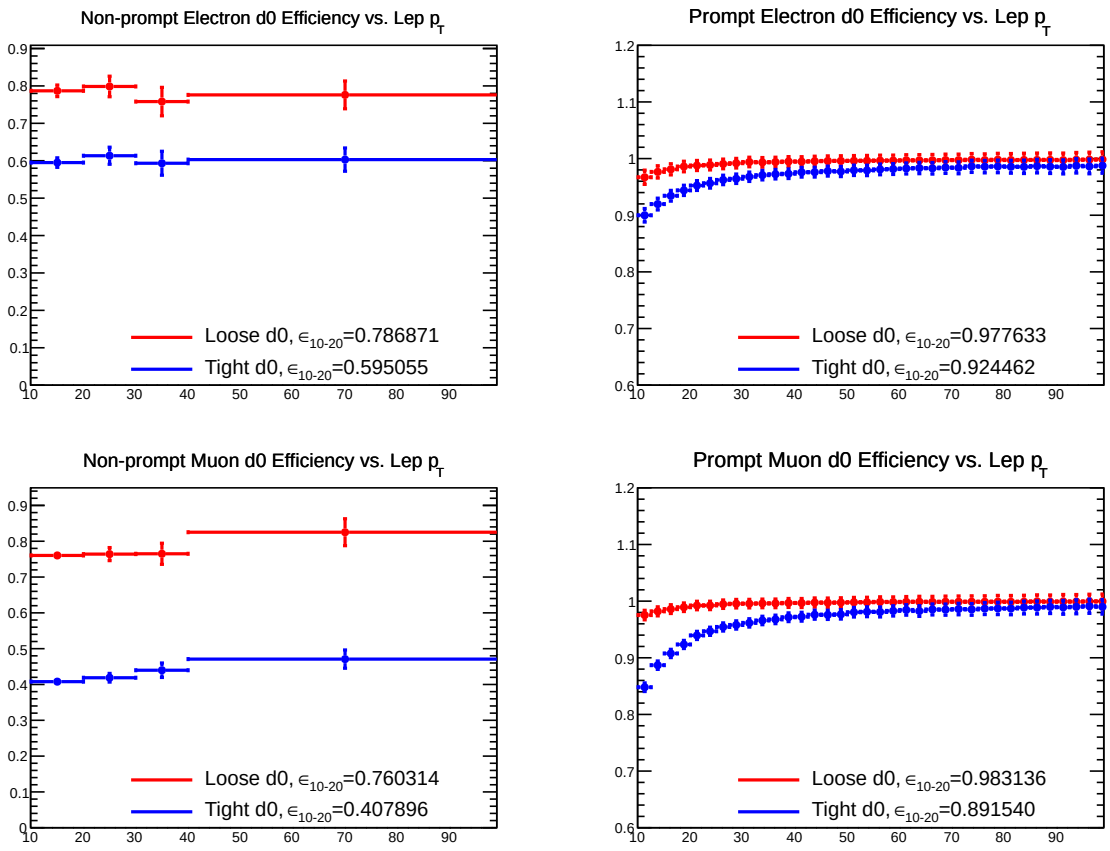


Figure 7-1. Single lepton efficiency of $|\mathrm{d}_{0,\mathrm{pv}}|$ is shown for loose and tight cuts.

7.2.3 B-Jet Selection

Identification of b-jets is a crucial part of this analysis. The lifetime of a b-quark is $\sim 10^{-12}$ seconds, and thus it can travel $\sim$ 0.03 cm from the primary vertex before decaying into a jet. This distance is enough to differentiate it from other jets with some reliability, and the jet can be fit to a secondary vertex for the event. The algorithm is called *CombinedSecondaryVertex*, based on secondary vertex and track impact parameter information is used to identify b-jets in this analysis, where the Medium Working point was deemed optimal [21, 22].

7.3 Trigger Strategy and Performance

The full list of the trigger names used in this analysis can be found in Tables 7-6 and 7-7. Triggers originally designed for Higgs analyses are used to select signal events

for both high-$\mathrm{p_T}$ and low-$\mathrm{p_T}$ analyses. Customized triggers were originally developed and commissioned to select signal events and control sample events specifically for this SUSY analysis. Though for the sake of combining analysis efforts for publication, their use was limited to control sample selection for the high-$\mathrm{p_T}$ and low-$\mathrm{p_T}$ analyses. These triggers are summarized in Tab. 7-7. These triggers rely on the signatures of dileptons+$\mathrm{H_T}$ as well as dileptons+$\not\mathrm{E}_\mathrm{T}$. The distinguishing feature of these triggers is that they do not impose any lepton isolation requirements online. This is a crucial feature that allows the nonprompt lepton backgrounds to be estimated with a large control sample, since the BTag&Probe method requires use of a sideband region with at least one lepton that has RelIso relaxed. If both leptons have isolation requirements at the trigger level selection, the sideband will be too small (or nonexistent) to use BTag&Probe reliably. The Higgs triggers have excellent signal efficiency, but they do impose isolation requirements online and are thus not useful for the nonprompt lepton background estimates.

Table 7-6. Triggers used for the high-$\mathrm{p_T}$ analysis signal and control sample selection.

Channel	Trigger names
$\mu\mu$	HLT_Mu17_Mu8
	HLT_DoubleMu8_Mass8
	HLT_DoubleMu14_Mass8
ee	HLT_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
	HLT_DoubleEle8_CaloIdT_TrkIdVL_Mass8
	HLT_DoubleEle14_Mass8_CaloIdT_TrkIdVL_pfMHT40
$\mathrm{e}\mu$	HLT_Mu17_Ele8_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
	HLT_Mu8_Ele17_CaloIdT_CaloIsoVL_TrkIdVL_TrkIsoVL
	HLT_Mu8_Ele8_CaloIdT_TrkIdVL_Mass8
	HLT_Mu14_Ele14_Mass8_CaloIdT_CaloIsoVL_TrkIdVL_pfMHT40

The Higgs triggers suffered some inefficiency at the beginning of 2012 data taking due to a software bug. Therefore efficiencies are measured separately for different run eras. Fig. 7-2 shows the trigger efficiencies for three different channels.

The efficiency of each component of the dilepton+$\mathrm{H_T}$ and dilepton+$\not\mathrm{E}_\mathrm{T}$ triggers are measured separately and then combined to get the total trigger efficiency. It is assumed

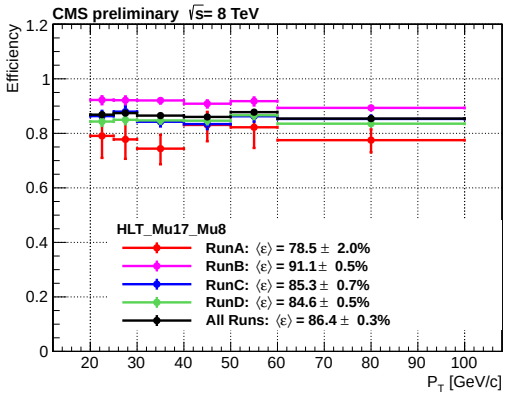
Table 7-7. Triggers used for the low-$\mathrm{p_T}$ analysis signal and control sample selection.

Channel	Trigger names
$\mu\mu$	HLT_DoubleMu8_Mass8_PFNoPUHT175
	HLT_DoubleMu8_Mass8_PFHT175
ee	HLT_DoubleEle8_CaloIdT_TrkIdVL_Mass8_PFNoPUHT175
	HLT_DoubleEle8_CaloIdT_TrkIdVL_Mass8_PFHT175
$\mathrm{e}\mu$	HLT_Mu8_Ele8_CaloIdT_TrkIdVL_Mass8_PFNoPUHT175
	HLT_Mu8_Ele8_CaloIdT_TrkIdVL_Mass8_PFHT175

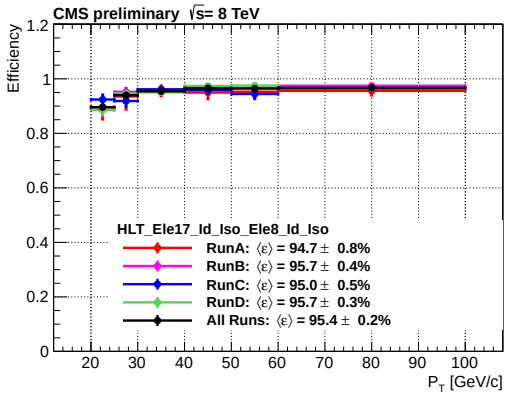
that the sources of inefficiency can be attributed to the differences in the online versions of the $\mathrm{H_T}$, $\not\mathrm{E}_\mathrm{T}$, and lepton $\mathrm{p_T}$ observables with respect to the offline versions, meaning differences in the HLT object reconstruction verses the full object reconstruction done offline. Further, it is assumed that these sources are mutually independent and thus can be factorized. Each part of the trigger is measured individually, taking the appropriate products to calculate the total efficiency for each trigger, respectively.

To measure the $\mathrm{H_T}$ and $\not\mathrm{E}_\mathrm{T}$ parts of the respective triggers we select events with the Higgs triggers so that the lepton part of the trigger becomes fully efficient. If the signal trigger fires, then the event is included in the numerator. The results are shown in Figs. 7-3 and 7-4. The apparent small inefficiency from the $\mathrm{H_T}$ component of the trigger strategy (Fig. 7-4) is actually due to the fact that the Higgs triggers feature slightly looser requirements on electrons. If there were inefficiencies due to the $\mathrm{H_T}$ part of the algorithm, it would be observed in the $\mu\mu$ channel, which it is not. Therefore the $\mathrm{H_T}$ component is taken to be 100% efficient for offline $\mathrm{H_T} > 200$ GeV in the high-$\mathrm{p_T}$ search and $\mathrm{H_T} > 250$ GeV in the low-$\mathrm{p_T}$ search . The $\not\mathrm{E}_\mathrm{T}$ efficiency is measured to range from 95% to 99% depending on the lepton channel.

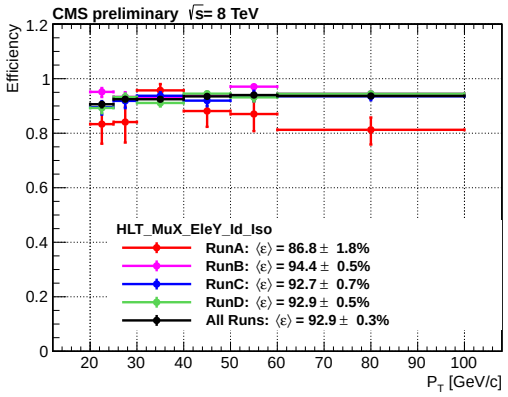
A similar strategy is applied for measuring the inefficiency of the lepton component, i.e. selecting events with hadronic triggers in the hadronic primary dataset, where the $\mathrm{H_T}$ or $\not\mathrm{E}_\mathrm{T}$ components are fully efficient, we obtain the unbiased efficiency dependence on the lepton $\mathrm{p_T}$ shown in Fig. 7-5. The efficiency is observed to be dependent on trailing lepton $|\eta|$ for the $\mu\mu$ channel and on trailing lepton $\mathrm{p_T}$ for the high-$\mathrm{p_T}$ ee



A



B



C

Figure 7-2. Trigger efficiency for signal events as a function of the softest lepton $\mathrm{p_T}$ for the (A) $\mu\mu$ (B) ee and (C) $\mathrm{e}\mu$ Higgs triggers.

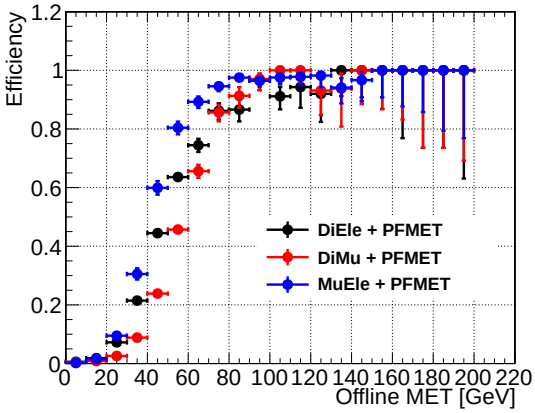


Figure 7-3. Trigger efficiency as a function of $\not\mathrm{E}_\mathrm{T}$.

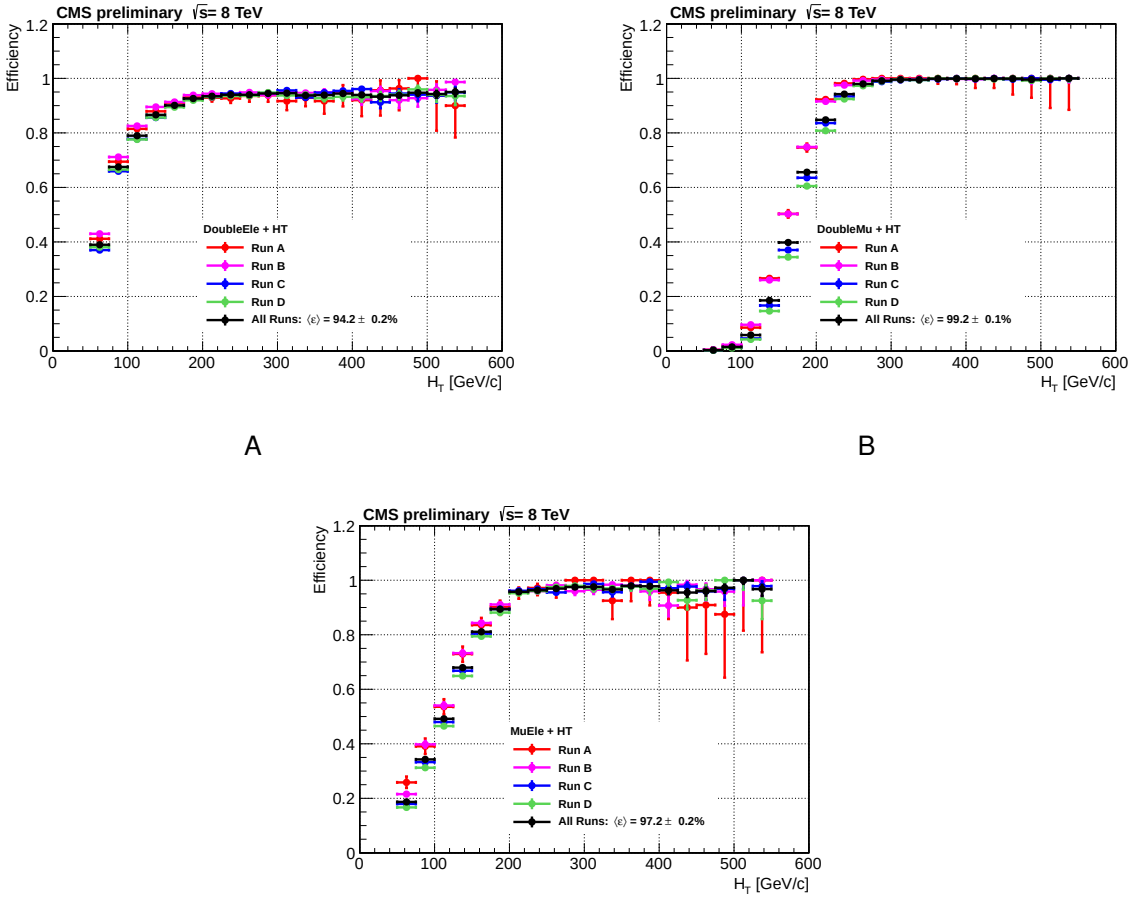


Figure 7-4. Trigger efficiency as a function of $\mathrm{H_T}$ for the (A) ee (B) $\mu\mu$ and (C) $\mathrm{e}\mu$ control sample triggers.

triggers. It is important to emphasize that these triggers are used for the nonprompt lepton background predictions which are assigned with systematic uncertainties of approximately 50%.

The measured Higgs trigger efficiencies will be used to scale down the Monte Carlo-based predictions of the irreducible backgrounds. These scale factors are also used in order to do hypothesis testing for potential signal models. Scale factors separated by channel for each analysis are shown in Table 7-8. An average systematic uncertainty of 6% will be applied across all channels to reflect the uncertainty in the trigger scaling that will be done to the Monte Carlo samples.

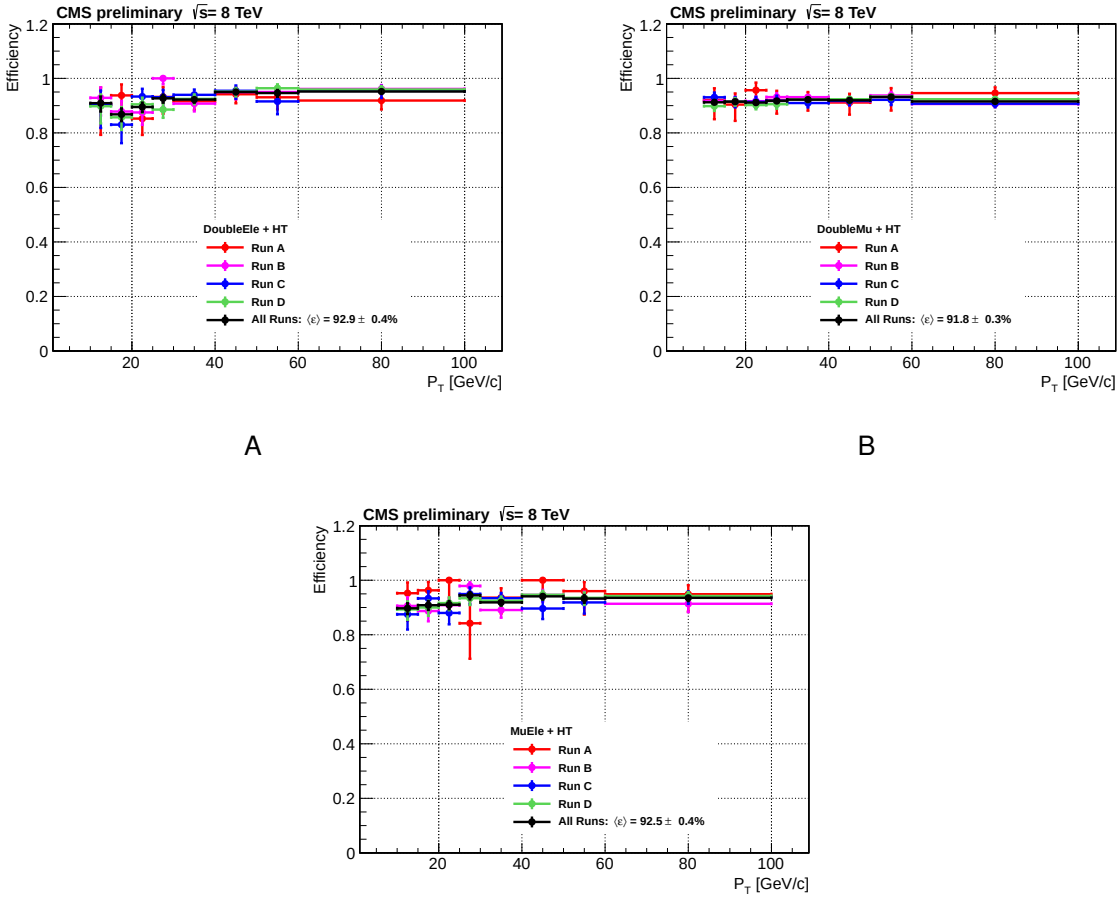


Figure 7-5. Trigger efficiency as a function of lepton $\mathrm{p_T}$ for the (A) ee (B) $\mu\mu$ and (C) $\mathrm{e}\mu$ control sample triggers.

7.4 Event Selection and Search Regions

The following selection requirements are imposed on the sample of events that pass the signal triggers listed in Tables 7-6 and 7-7:

- At least one good primary vertex. In the case of multiple vertices, the one with the hardest $\Sigma\mathrm{p_T}$ is chosen as the event vertex.
- At least one pair of same-sign leptons, $\mu\mu$, $\mathrm{e}\mu$ or ee where muons and electrons pass the requirements listed in Tables 7-1 and 7-2 respectively.
- The two candidate leptons must not combine to form an invariant mass below 8 GeV. This selection is applied in order to suppress events from heavy flavor quark decays.

Table 7-8. Trigger scale factors from trigger inefficiencies applied to Monte Carlo predictions of the *irreducible* backgrounds. For some channels, scale factors are parametrized by trailing lepton $\mathrm{p_T}$ or $|\eta|$.

low-$\mathrm{p_T}$	Scale Factor	high-$\mathrm{p_T}$	Scale Factor
$\mu\mu$, $\|\eta\| < 1$	0.94	$\mu\mu$, $\|\eta\| < 1$	0.90
$\mu\mu$, $\|\eta\| > 1$	0.90	$\mu\mu$, $\|\eta\| > 1$	0.81
$\mathrm{e}\mu$	0.93	$\mathrm{e}\mu$	0.93
ee	0.93	ee, $\mathrm{p_T} < 30$	0.92
		ee, $\mathrm{p_T} > 30$	0.96

- Veto signal leptons which pair with other leptons, which are selected using the loose selection criteria shown in Tables 7-3 and 7-4, of the same flavor and opposite charge, and form an invariant mass consistent with a Z-decay within ± 15 GeV from the Z mass.
- Extend the Z-veto to veto γ^* events if a signal lepton forms an opposite-sign, same-flavor pair with a loose lepton having invariant mass < 12 GeV.
- For events that may have multiple pairs of SS leptons, the preference is given by channel: $\mu\mu$, $\mathrm{e}\mu$, and ee. Within a given channel the preference is given to the pair with the highest $\Sigma|\mathrm{p_T^{\ell}}|$.
- Finally, require events to have sufficient hadronic activity in order to obtain a sample of events that pass the trigger requirement with high efficiency. This requirement comprises of number of jets, $\mathrm{N_{jets}} \geq 2$ and $\mathrm{H_T} > 200$ GeV.
- $\not\mathrm{E}_\mathrm{T} > 50$ GeV

Using events that pass the above described criteria multiple search regions are defined in Fig. 7-6, which are motivated by a variety of SUSY scenarios. In processes such as $\mathrm{pp} \rightarrow \tilde{\mathrm{g}}\tilde{\mathrm{g}} \rightarrow \mathrm{tttt}\tilde{\chi}^0\tilde{\chi}^0$ and $\mathrm{pp} \rightarrow \tilde{\mathrm{b}}\tilde{\mathrm{b}} \rightarrow \mathrm{ttWW}\tilde{\chi}^0\tilde{\chi}^0$ high jet multiplicity is expected. But for light stops and sbottoms, these jets can be soft and fail the selection, so there are search regions with 2-3 jets to gain sensitivity to this scenario. In the second process above, two b-jets are expected, but in case one fails the selection, there is still sensitivity to this process in the $\mathrm{N_{b-jets}} = 1$ search regions. In addition, there can be small mass splittings between SUSY particles which would result in low $\not\mathrm{E}_\mathrm{T}$, so there is a $\not\mathrm{E}_\mathrm{T}$ bin for 50-120 GeV. The high-$\mathrm{p_T}$ search is ideal for on-shell W's in these decays,

but off-shell W's can lead to low-p_T leptons, which is why 10 GeV leptons are included for the low-p_T analysis.

N_{b-jets}	$\not{E}_T$	N_{jets}	H_T $[200-400]$	$H_T > 400$
$= 0$	$50-120$	$2-3$	SR01	SR02
		≥ 4	SR03	SR04
	> 120	$2-3$	SR05	SR06
		≥ 4	SR07	SR08
$= 1$	$50-120$	$2-3$	SR11	SR12
		≥ 4	SR13	SR14
	> 120	$2-3$	SR15	SR16
		≥ 4	SR17	SR18
≥ 2	$50-120$	$2-3$	SR21	SR22
		≥ 4	SR23	SR24
	> 120	$2-3$	SR25	SR26
		≥ 4	SR27	SR28

Figure 7-6. Signal region labels and definitions high-p_T and low-p_T analyses. The high-p_T search regions have an H_T threshold of 200 GeV, as shown here, while the low-p_T analysis have $H_T > 250$ GeV, where the low-p_T triggers are fully efficient.

In order to test the background prediction methods, four baseline regions are used. These baselines are defined with a loose event selection in order to encompass the search regions while still showing the kinematic regions of interest. The baselines are defined in Fig. 7-7.

Analysis	BR	$\not{E}_T$	N_{b-jets}	H_T
High-p_T	H0	> 30	≥ 0	> 200
	H2	> 30	≥ 2	> 200
Low-p_T	L0	> 30	≥ 0	> 250
	L2	> 30	≥ 2	> 250

Figure 7-7. Baseline regions (BR) for the high-p_T and low-p_T searches where $N_{b-jets} \geq 0, 2$.

7.5 Background Predictions

In this section, the observed yields in each of the chosen search regions are reported, as well as the Standard Model background predictions. The categorization of

the main backgrounds, as well as the methods used to estimate them, are summarized in Table 7-9.

Table 7-9. Classification of background processes and the methods used to predict each.

Background Type	Sources	Method
Same-sign prompt-prompt (N^{SS}_{p-p})	$t\bar{t}W$, $t\bar{t}Z$, WZ, ZZ,	Monte Carlo-based
Opposite-sign prompt-prompt (N^{OS}_{p-p})	Charge-flip in $t\bar{t}$, tW, DY, $W^{\mp}$, etc	Data-driven (Charge-flip method)
Same-sign prompt-nonprompt (N^{SS}_{p-n})	$t\bar{t}$, tW, W+jets, Z+jets	Data-driven (BTag&Probe Method)
Same-sign nonprompt-nonprompt (N^{SS}_{n-n})	QCD, all-hadronic $t\bar{t}$	Data-driven (BTag&Probe Method)

For each search region the total background prediction is given by 7–3.

$$N^{tot}_{bgd} = N^{SS}_{p-p} + N^{OS}_{p-p} + N^{SS}_{n-n} + N^{SS}_{p-n} \tag{7–3}$$

Before proceeding with the evaluation of backgrounds, it is convenient to organize the regions of interest that will be used in this analysis. Table 7-10 summarizes the purpose and event selection criteria for each one. The regions $b\bar{b}$ Control 1 and $b\bar{b}$ Control 2 are used to measure the *fake rate* of nonprompt leptons, i.e. the probability with which nonprompt leptons are isolated and therefore fake the signature of a prompt lepton. There is a Sideband region for every search region, and it contains the number of events to which the fake rate is applied, thus providing N^{SS}_{n-n} and N^{SS}_{p-n} from Eq. 7–3. The regions Z Control 1 and Z Control 2 are used for determining the rate of electron charge misidentification, while the region OS Control provides the number of events to which this charge flip rate is applied, providing the N^{OS}_{p-p} piece of the background. The N^{SS}_{p-p} background is the irreducible background obtained from Monte Carlo generated events.

Table 7-10. Control regions used for background estimation.

Region	Requirements	Comments
Sideband	$\ell^{\pm}\ell^{\pm}$ (inclusive) Min$\{\mathrm{RelIso}(\ell_1), \mathrm{RelIso}(\ell_2)\} < 0.1(0.09)$ Max$\{\mathrm{RelIso}(\ell_1), \mathrm{RelIso}(\ell_2)\} > 0.1(0.09)$ $\mathrm{H_T} > 200/320$ GeV $\not\!\!\mathrm{E}_{\mathrm{T}} > 50/120$ GeV	Predicted yields from single-t and $\mathrm{t\bar{t}}$ production in the signal regions will be extrapolated from these regions via the BTag&Probe Method
$\mathrm{b\bar{b}}$ Control 1	one or two b-tagged jets w/ $\mathrm{p_T} > 40$ GeV $\ell^{\pm}$ w/ $\mathrm{p_T} > 10$ GeV $\mathrm{RelIso}(\ell)$ relaxed $\mathrm{M_T} < 15$ GeV $\not\!\!\mathrm{E}_{\mathrm{T}} < 15$ GeV	Measuring the selection efficiencies for $\mathrm{RelIso}(\mu) < 0.1$ and $\mathrm{RelIso(e)} < 0.09$ using the BTag&Probe Method to predict prompt-fake background in $\mathrm{N_{b-jets}}$= 0,1 region
$\mathrm{b\bar{b}}$ Control 2	two b-tagged jets w/ $\mathrm{p_T} > 40$ GeV $\ell^{\pm}$ w/ $\mathrm{p_T} > 10$ GeV $\mathrm{RelIso}(\ell)$ relaxed $\mathrm{M_T} < 15$ GeV $\not\!\!\mathrm{E}_{\mathrm{T}} < 15$ GeV	Measuring the selection efficiencies for $\mathrm{RelIso}(\mu) < 0.1$ and $\mathrm{RelIso(e)} < 0.09$ using the BTag&Probe Method to predict prompt-fake background in $\mathrm{N_{b-jets}} \geq$ 2 region
Z Control 1	$\mathrm{RelIso}(\ell_1) < 0.15$ $76 < \mathrm{M}_{\ell_1\ell_2} < 106$ $\mathrm{H_T}$ relaxed, $\not\!\!\mathrm{E}_{\mathrm{T}} < 30$ GeV	Measuring the $\mathrm{e^{\pm}e^{\mp}}$ yields for the Charge Flip Method
Z Control 2	Same as Z Control 1 but with $\ell^{\pm}\ell^{\pm}$ (exclusive)	Measuring the $\mathrm{e^{\pm}e^{\pm}}$ yields for the Charge Flip Method
OS Control	Same as Signal regions but with opposite-sign leptons	Measuring the $\mathrm{e^{\pm}e^{\pm}}$ yields for the Charge Flip Method

7.5.1 Prompt-Nonprompt Background Prediction

In order to predict the backgrounds due to leptons coming from jets (mainly from $\mathrm{t\bar{t}}$ production), a data-driven technique is used that was originally developed for an early version of the SS dilepton analysis where the main source of the background is due to leptons coming from b-jets. It is known as the BTag&Probe method and has been documented in detail previously [23–26].

The estimation of the number of the events with a nonprompt lepton is based on the extrapolation from the isolation sideband for each signal region. An isolation sideband is selected in the same way as a signal region with the only change that the isolation requirement on one of the leptons (a probe lepton) is inverted: $\mathrm{RelIso}(\mu(\mathrm{e})) > 0.09(0.10)$. All the other steps of the selection in the sideband are carried out with the probe lepton

being treated as a signal lepton, i.e. the selected jets are cleaned with respect to this lepton, both Z- and γ^*-vetos are applied, and so on. This approach ensures that all the kinematic distributions of interest ($\mathrm{H_T}$, $\not\mathrm{E_T}$, $\mathrm{N_{jets}}$, $\mathrm{N_{b-jets}}$) are correctly described.

The control sample is selected to reproduce effects leading to the appearance of an isolated lepton coming from an untagged or tagged b-jet as well as to incorporate some fraction of misidentified leptons from light jets. Ideally, one would like to select a control sample that is as close as possible to the signal selection. In order to achieve this, two $\mathrm{b\bar{b}}$ control samples, 1 and 2, are defined. The ratio of probabilities for a lepton to be isolated and non-isolated (the fake rate, FR) is measured and used in for the extrapolation from a sideband. The efficiencies measured in $\mathrm{b\bar{b}}$ Control 1 are used to predict the background for $\mathrm{N_{b-jets}} = 0, 1$ search regions, and the efficiencies measured in $\mathrm{b\bar{b}}$ Control 2 for predicting the nonprompt lepton background search regions where two or more b-tagged jets are required.

In the case of $\mathrm{N_{b-jets}} = 0, 1$ in the event, one of the two leptons comes from a W decay and another from a semi-leptonic decay of a b-quark. When one moves to two or more b-jets in the event, the second lepton has to be either a misidentified light jet or a product of a decay of a b-quark which does not come from a top-quark. The latter two processes are also present in the case of $\mathrm{N_{b-jets}} = 0, 1$, but they constitute a small fraction compared to the main source of nonprompt leptons, that is b-quarks from top decay. Since the source of the nonprompt leptons is a $\mathrm{b\bar{b}}$-pair associated with a $\mathrm{t\bar{t}}$, the mother partons of such leptons do not have necessarily the same hard spectrum as the b-jets from $\mathrm{t\bar{t}}$. Therefore, the fake rate can behave differently.

Available MC allows a view of the claimed difference, though only one sample ($\mathrm{t\bar{t}}$@NLO) was used. Other samples exhibit unexpected behavior: b-quarks almost never have a top-quark among their mothers. The closest jet $\mathrm{p_T}$ spectra for two types of b-quarks (from top or not from top) are shown in Fig. 7-8.

Here is the summary of the definition of these control samples:

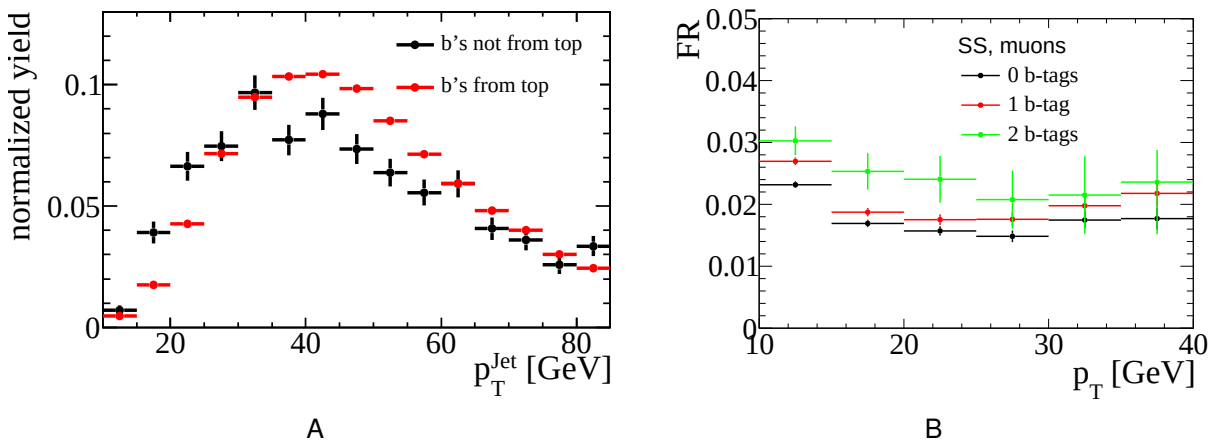


Figure 7-8. Muon fake rate for $\mathrm{t\bar{t}}$ events passing SS dilepton selection and having a nonprompt muon: (A) the closest jet $\mathrm{p_T}$ for nonprompt muons with $\mathrm{p_T} < 20$ GeV and (B) the fake rate for muons.

- $\mathrm{b\bar{b}}$ Control 1:
  - at least one electron/muon that passes all the lepton ID requirements (except the impact parameter and isolation requirement) with $\mathrm{p_T}(\ell) > 10$ GeV.
  - $\mathrm{M_T} < 20$ GeV, suppress the W/Z contamination.
  - $\not\mathrm{E}_\mathrm{T} < 20$ GeV, suppress the W/Z contamination.
  - $\mathrm{d_0} > 0.005$ for leptons with $\mathrm{p_T} > 25$ GeV to suppress contamination with prompt leptons from $\mathrm{t\bar{t}}$ and $\mathrm{W^{\pm}}$+jets.
  - b-jets coming from gluon-splitting are typically soft and their $\mathrm{p_T}$ spectrum is very different from b-jets in $\mathrm{t\bar{t}}$. Typically, in the case of gluon splitting, the azimuthal angle between the two b-jets is small, so their contribution can be diminished by requiring $\Delta\phi(\mathrm{jet}^1,\ \mathrm{jet}^2) > 2$ for dijet events.
  - At least 1 jet that passes the b-jet selection described in Section 7.2.3.
  - Only leptons with $\mathrm{p_T} \leq 35$ GeV are considered for the efficiency templates. This is done in order to reduce contamination from prompt lepton sources (e.g. $\mathrm{t\bar{t}}$), which is especially important in events with higher jet multiplicities. Implicit in this choice is the assumption that beyond $\mathrm{p_T} = 35$ GeV, the nonprompt lepton isolation efficiency is flat as a function of $\mathrm{p_T}$.
- $\mathrm{b\bar{b}}$ Control 2:
  - same selection as for $\mathrm{b\bar{b}}$ Control 1 but here at least two b-jets are required in the event.

Using the above described control samples 1 and 2, the calculation of the nonprompt isolation efficiencies is performed, to be used in search regions with $\mathrm{N_{b-jets}} = 0, 1$ required and the search regions with $\mathrm{N_{b-jets}} \geq 2$ required, respectively. The isolation of the nonprompt leptons are measured in bins of lepton-$\mathrm{p_T}$.

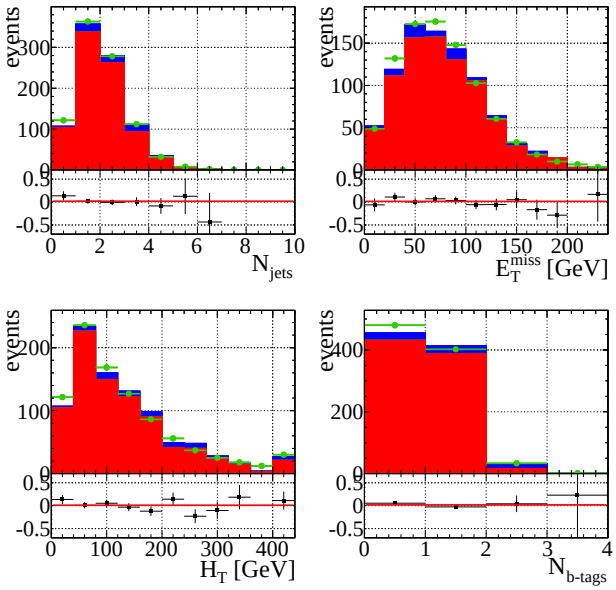
The performance of the method is tested in $\mathrm{t\bar{t}}$ MC samples with the fake rate measured in $\mathrm{b\bar{b}}$ events. The closure test is performed in the selection requiring only two SS leptons in $\mathrm{t\bar{t}}$, but it is checked as a function of the variables used to define the search regions in the analysis. The results are shown in Fig. 7-9.

In the search regions with b-tags, W+jets can also play non-negligible role, therefore the closure test is carried out for this process as well. Since the nonprompt leptons come from softer jets in this case, it is expected that the FR for W+jets is higher in general. The spectrum of these jets is directly related to the $\mathrm{H_T}$ measured in the event (while in $\mathrm{t\bar{t}}$, $\mathrm{H_T}$ is impacted also by the number of W which decayed hadronically). Since the baseline selection requires $\mathrm{H_T} > 250(200)$ GeV and $\mathrm{N_{jets}} \geq 2$, it models the jet spectrum in the W+jets events selecting the hard part of it. In Figures 7-10 and 7-11, the closure test is shown for W+jets samples without additional requirements and with requiring $\mathrm{H_T} > 160$ GeV, $\mathrm{N_{jets}} \geq 2$. The selection is done by requiring 2 leptons (same-sign or opposite-sign) fully satisfying lepton selection criteria but with the $\mathrm{d_0}$ cut relaxed.

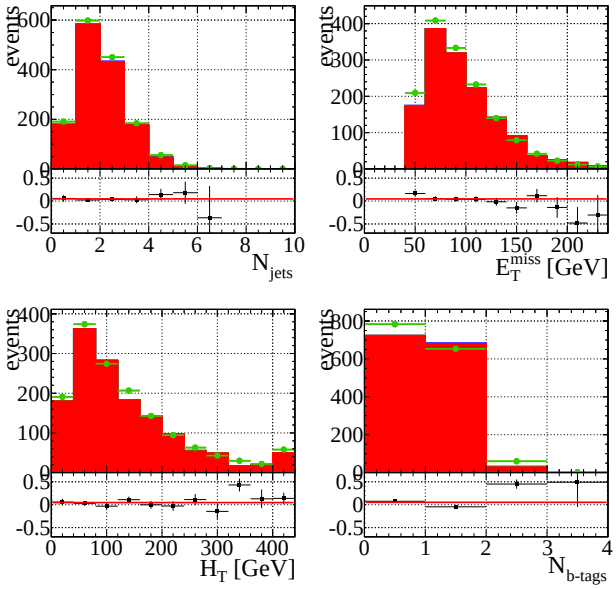
The MC samples used for the closure test in W+jets are listed in Table 7-11. These samples are added together without a cross-section weight applied.

Table 7-11. W+jets MC samples. All samples are produced with the generator.

DBS Name	σ (pb)
/WJetsToLNu_HT-200To250_8TeV-madgraph/Summer12_DR53X-PU_S10_START53_V7C-v1	90.27
/WJetsToLNu_HT-250To300_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	48.01
/WJetsToLNu_HT-300To400_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	38.3
/WJetsToLNu_HT-400ToInf_8TeV-madgraph_v2/Summer12_DR53X-PU_S10_START53_V7A-v1	25.22

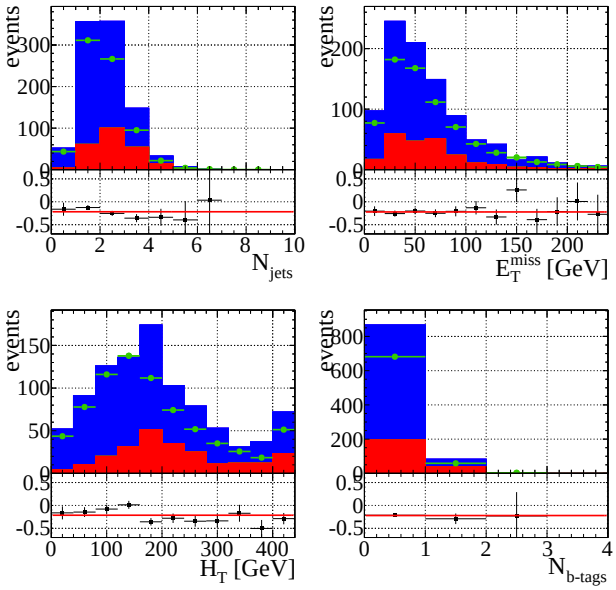


A

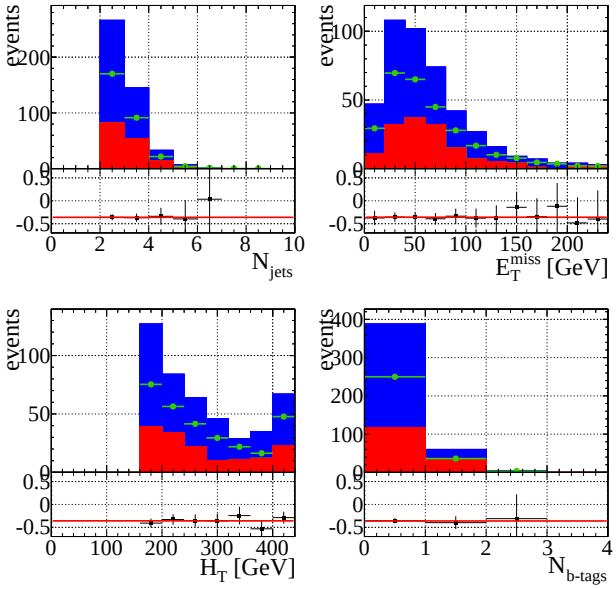


B

Figure 7-9. Closure test in $\mathrm{t\bar{t}}$ MC with SS selection for (A) nonprompt e and (B) nonprompt μ. Observed number of events is shown by filled histograms: red shows events where a nonprompt lepton comes from a b-jet, green shows leptons from c-jet, blue shows leptons from light jets. Predicted number of events is shown by points. The red line corresponds to a fit with a constant.

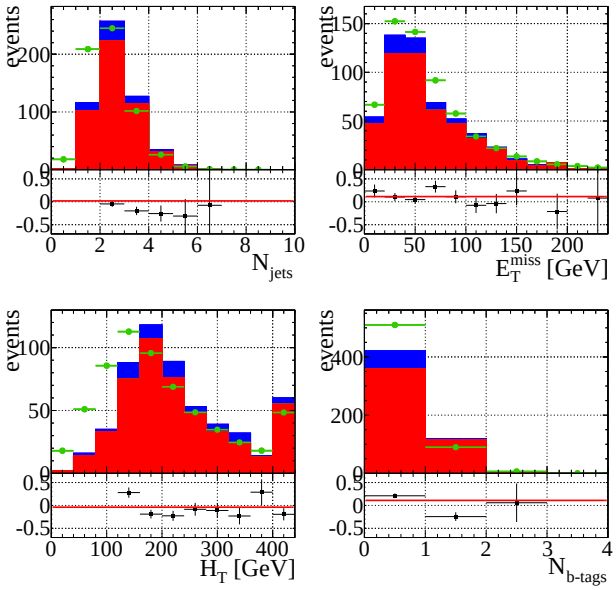


A

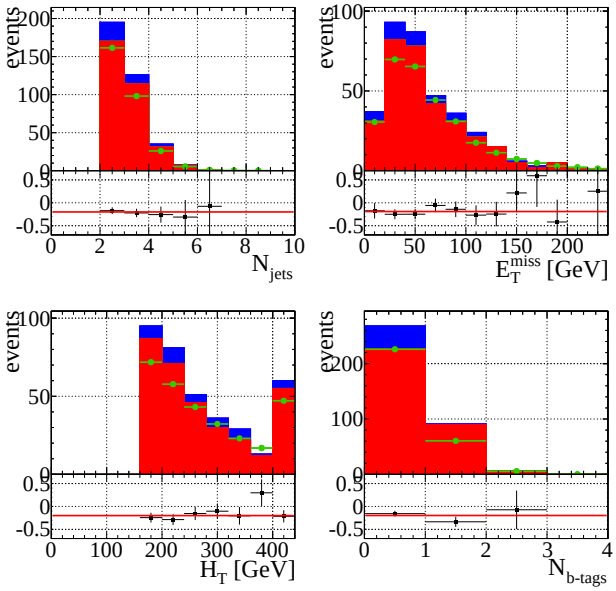


B

Figure 7-10. Closure test for electrons in W+jets MC with dilepton selection for (A) nonprompt e and (B) $\mathrm{H_T} > 160$ GeV, $\mathrm{N_{jets}} \geq 2$. Observed number of events is shown by filled histograms: red shows events where a nonprompt lepton comes from a b-jet, green shows leptons from c-jet, blue shows leptons from light jets. Predicted number of events is shown by points. The red line corresponds to a fit with a constant.



A



B

Figure 7-11. Closure test for muons in W+jets MC with dilepton selection for (A) nonprompt μ and (B) $\mathrm{H_T} > 160$ GeV, $\mathrm{N_{jets}} \geq 2$. Observed number of events is shown by filled histograms: red shows events where a nonprompt lepton comes from a b-jet, green shows leptons from c-jet, blue shows leptons from light jets. Predicted number of events is shown by points. The red line corresponds to a fit with a constant.

When measuring FR in $b\bar{b}$ MC, the cuts on $\not\!\!E_T$ and M_T are not applied to increase the available statistics. The influence of the cut on $\not\!\!E_T$ is studied separately and the result is shown in Figure 7-12.

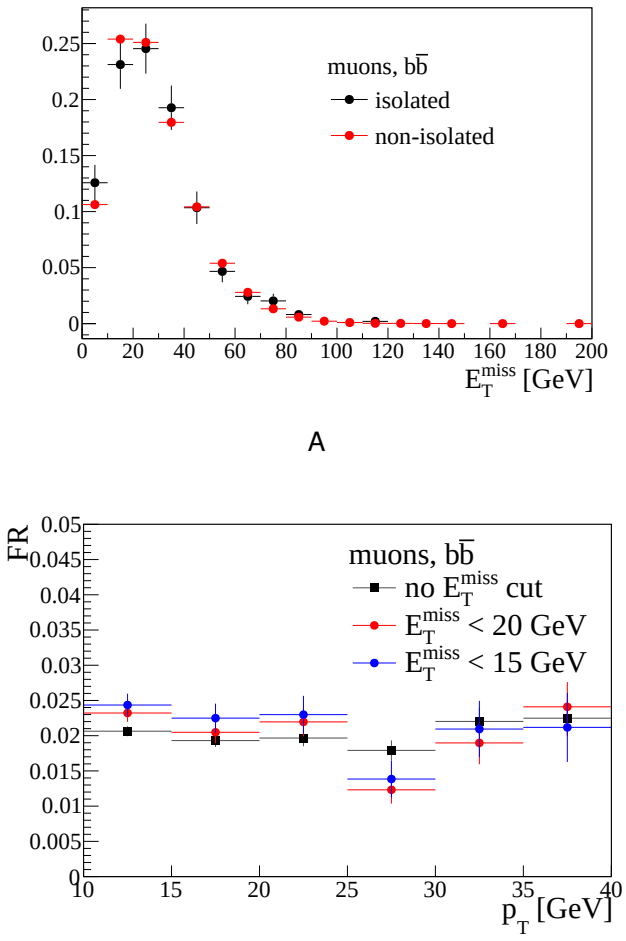
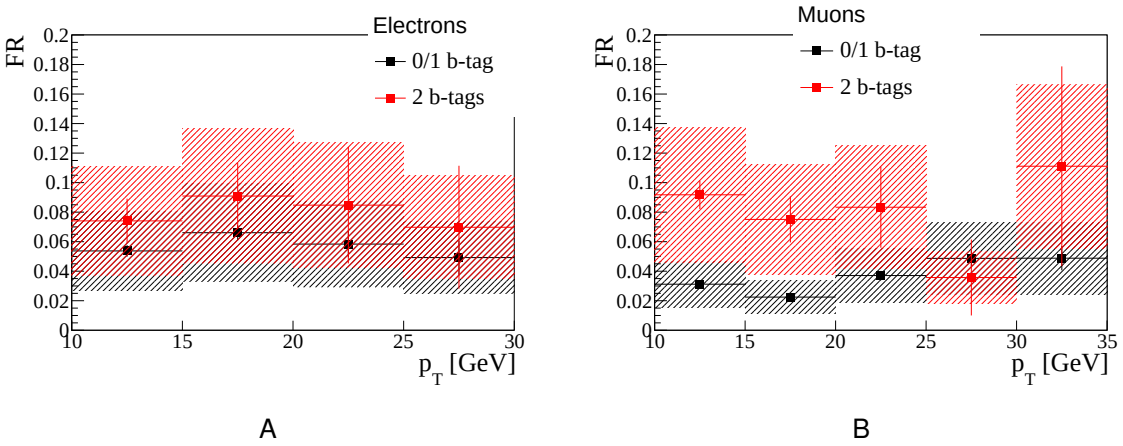


Figure 7-12. Fake rate as a function of $\not\!\!E_T$. (A) The $\not\!\!E_T$ spectrum in $b\bar{b}$ events with isolated and non-isolated muons. (B) The FR measured with various cuts on $\not\!\!E_T$.

The systematic uncertainties associated with measurements coming from the BTag&Probe method have been carefully assessed in the analysis [23]. These uncertainties continue to be appropriate (and conservative), and are again used here. To reiterate, the main sources of the systematic uncertainty are the closure test, the prompt lepton contamination in the control samples, and the possible mis-match of the sources

of the nonprompt leptons in the the control sample and the background. The total systematic uncertainty for electron(muon) efficiencies is 50(50)%. These uncertainties are propagated to the final prediction assuming 100% correlation.

The fake rates measured in control samples in data are summarized in Fig. 7-13. The FR for $\mathrm{N_{b-jets}} \geq 2$ search regions is found to be higher than that of $\mathrm{N_{b-jets}} = 0, 1$. This fact is consistent with the MC studies and corresponds to having softer spectrum of lepton mothers in the 2 b-jet scenario.

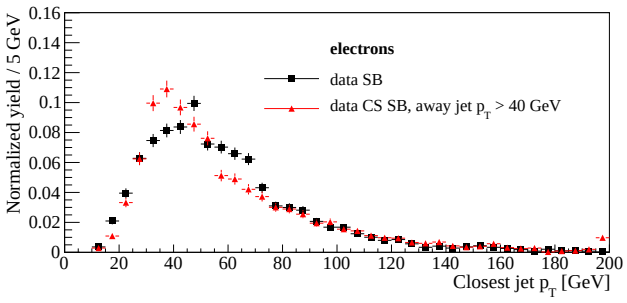


A B

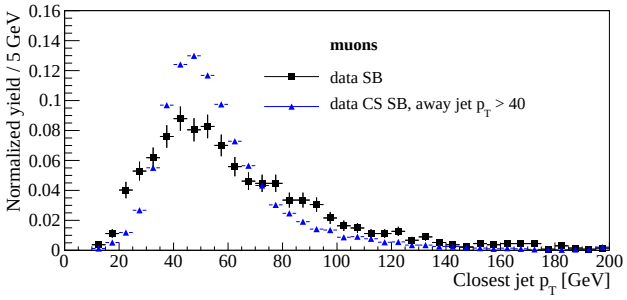
Figure 7-13. Fake rate from MC: (A) for electrons and (B) for muons. The filled area shows the systematic uncertainty.

The $\mathrm{p_T}$ spectra of the jets containing a non-isolated electron or muon for both the sideband in data and the data control samples are shown in Fig. 7-14. They show reasonable agreement and justify the selection which is applied to the control sample. These plots contain closest jet $\mathrm{p_T}$ spectra without a differentiation in number of b-tags. The $\mathrm{N_{b-jets}} \geq 2$ region does not have enough statistics to draw any conclusions about the shape of the distribution.

The dependence on the mother parton spectrum can be reduced by redefining the sideband region. The entries in the tail of the isolation distribution are formed with soft leptons residing within a hard jet. With resorting to the short sideband (e.g. with the requirement that a sideband lepton should have $\mathrm{RelIso} < 1$), the hard part of the



A



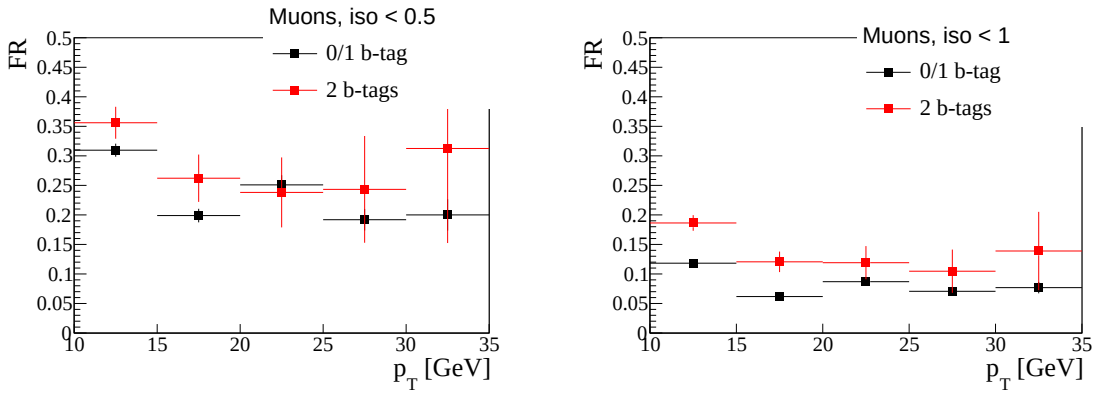
B

Figure 7-14. Closest jet $\mathrm{p_T}$ of the jet containing a non-isolated lepton in a sideband used to estimate nonprompt background and in the control sample for (A) electrons and (B) muons.

mother-jet $\mathrm{p_T}$ spectrum is significantly suppressed and therefore the fake rates for 0/1 and 2 b-jets scenarios are more similar. This is confirmed by the measurement of the fake rate in data with two variations of the short sideband: $0.09(0.10) < \mathrm{RelIso} < 0.50$ (Fig. 7-15A) and $0.09(0.10) < \mathrm{RelIso} < 1.00$ (Fig. 7-15B).

For a consistency check, the results for the SS analysis selection in one of the baseline regions were obtained by using the default (large) sideband and by using a short one ($\mathrm{RelIso} < 0.5$). The baseline selection is defined with $\mathrm{H_T} > 200$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV and $\mathrm{N_{jets}} > 1$. The results are shown in Table 7-12. The corresponding kinematic distributions can be found in Fig. 7-16 and 7-17. The two variations of the method demonstrate consistent results and can serve as a robustness check for the BTag&Probe results.

The background due to two nonprompt leptons, namely from QCD-multijet production, are estimated using the sideband selected with two non-isolated leptons.



A B

Figure 7-15. Fake rate from data: (A) for electrons and (B) for muons. The filled area shows the systematic uncertainty.

Table 7-12. Baseline RelIso consistency check. Baseline defined with SS dileptons with $\mathrm{p_T} >$ 10 GeV, $\mathrm{N_{jets}} > 1$, $\mathrm{H_T} > 200$ GeV, $\not{\mathrm{E}}_\mathrm{T} > 30$ GeV, $\mathrm{N_{b-jets}} \geq 0$. The number of expected events with at least one nonprompt lepton is evaluated by using a long (inverted RelIso) and a short ($\mathrm{RelIso} < 0.5$) sidebands.

$\mathrm{N_{b-jets}}$	long SB	short SB
≥ 0	217	234
1	98	112
≥ 2	22	23

The nonprompt lepton isolation efficiencies are used, as measured in the BTag&Probe method. The reason for this is to avoid possible bias in the isolation efficiency measurement due to prompt lepton contamination.

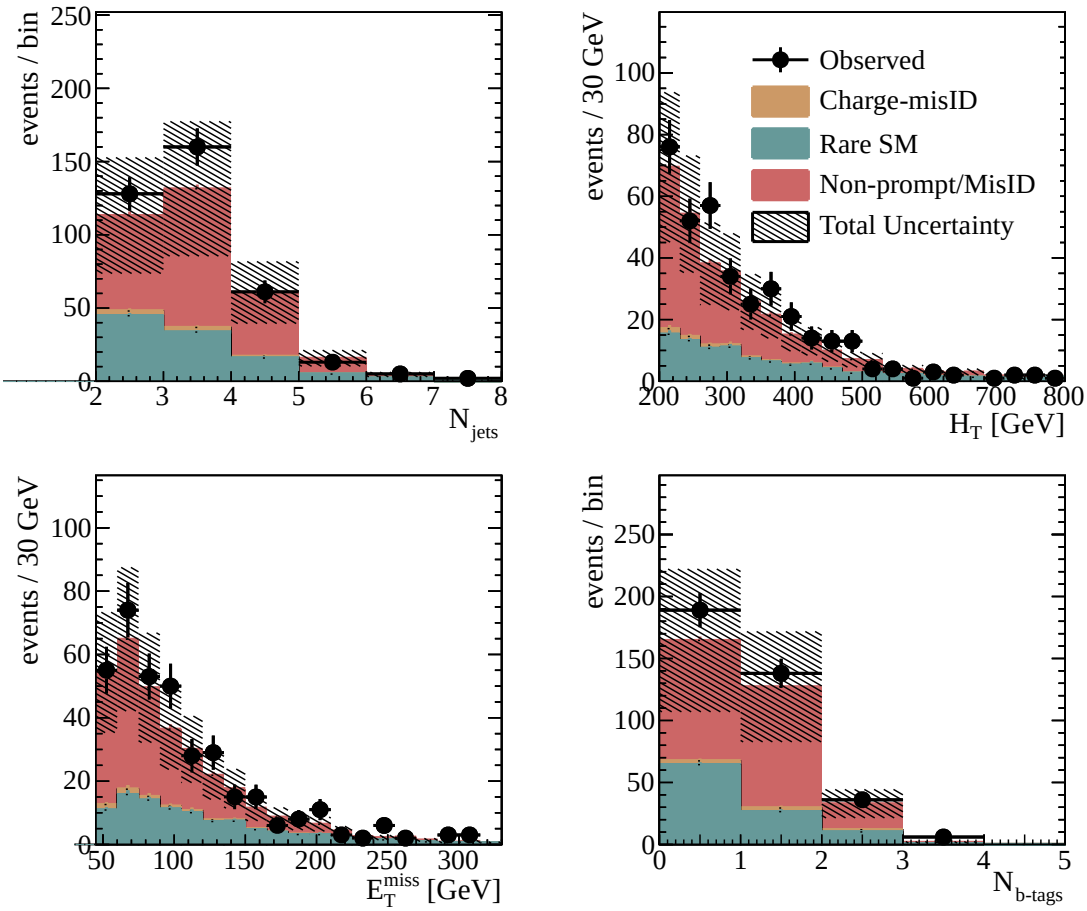


Figure 7-16. Kinematic distributions using the long sideband for SS dilepton selection with the requirements $\mathrm{H_T} > 200$ GeV, $\not\!\!\mathrm{E}_\mathrm{T} > 30$ GeV and $\mathrm{N_{jets}} > 1$.

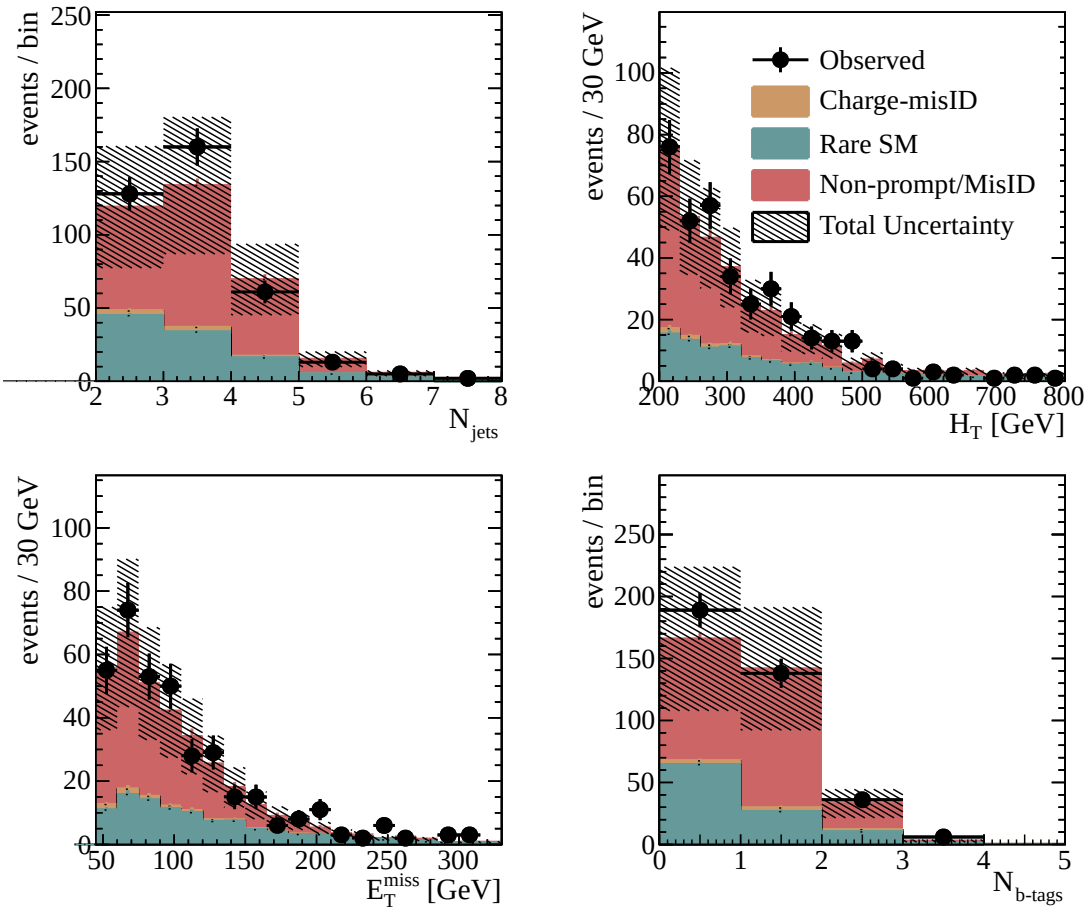


Figure 7-17. Kinematic distributions using the short sideband for SS dilepton selection with the requirements $\mathrm{H_T} > 200$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV and $\mathrm{N_{jets}} > 1$.

Table 7-13. Charge mis-id Rate for loose $|\mathrm{d}_{0,\mathrm{pv}}|$.

η Region	Charge-flip rate in Data	Charge-flip rate in MC
Barrel-Barrel	$(1.7 \pm 0.2) \times 10^{-4}$	$(1.4 \pm 0.1) \times 10^{-4}$
Barrel-Endcap	$(7.0 \pm 0.4) \times 10^{-4}$	$(5.3 \pm 0.3) \times 10^{-4}$
Endcap-Endcap	$(19.5 \pm 1.8) \times 10^{-4}$	$(13.6 \pm 1.3) \times 10^{-4}$

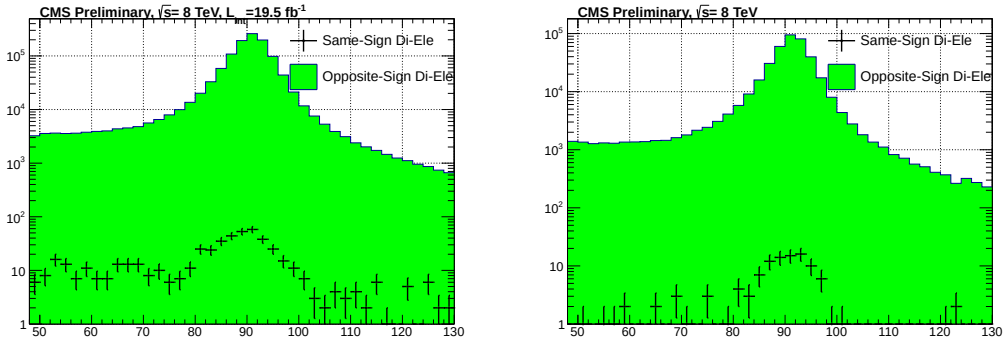
Table 7-14. Charge mis-id Rate for tightened $|\mathrm{d}_{0,\mathrm{pv}}|$.

η Region	Charge-flip rate in Data	Charge-flip rate in MC
Barrel-Barrel	$(0.65 \pm 0.07) \times 10^{-4}$	$(0.61 \pm 0.08) \times 10^{-4}$
Barrel-Endcap	$(3.0 \pm 0.2) \times 10^{-4}$	$(1.6 \pm 0.2) \times 10^{-4}$
Endcap-Endcap	$(6.4 \pm 0.8) \times 10^{-4}$	$(4.9 \pm 0.7) \times 10^{-4}$

7.5.2 Prediction Charge Mis-ID Background

The probability for electron charge to be mis-measured is non-negligible and contributes to a sub-leading background to the same-sign final state topology. A measurement of the charge-flip probability $\varepsilon_{\text{flip}}$ was performed by measuring the number of SS electron pairs under the Z-peak ($\pm$ 15 GeV) at $\sqrt{\mathrm{s}} = 8$ TeV. The ratio of the number of events with SS dileptons to the OS dileptons, forming an invariant mass close to the Z mass, is used to calculate the charge mis-id rate, shown in Table 7-14.

This measurement was made with the full $\sqrt{\mathrm{s}} = 8$ TeV dataset and with loose and tight $|\mathrm{d}_{0,\mathrm{pv}}|$ for electrons and is shown in Table 7-13 and Table 7-14, respectively. The charge mis-id rate was reduced by a factor of 2-3 when $|\mathrm{d}_{0,\mathrm{pv}}|$ was tightened. Figure 7-18 shows the di-electron mass distributions that are used. Charge mis-measurement rate for electrons differ significantly as a function of pseudo-rapidity due to inhomogeneity of the material budget in the detector. Therefore, the charge mis-id rate is measured separately for the cases when both electrons are detected in the barrel region of the ECAL, one is in the barrel and one is in the endcap and for when both are in the endcap. The measured charge mis-id rates are used together with the event yields in OS Control region to obtain the final number of backgrounds due to charge mis-measurement.



A B

Figure 7-18. Invariant mass distributions used for charge mis-id rate in (A) data and (B) MC. The ratio of the total SS yield to OS yield is used to calculate the charge mis-measurement probability for Data and MC.

7.5.3 Same-Sign Prompt-Prompt background prediction

Now that the LHC has produced several femtobarns of data, this analysis is vulnerable to extremely rare Standard Model processes, which have yet to be measured directly (e.g., $\mathrm{W^{\pm}W^{\pm}}$, $\mathrm{t\bar{t}W^{\pm}}$, $\mathrm{t\bar{t}Z}$, $\mathrm{t\bar{t}\gamma^{*}}$, WWZ, WWW, WZZ, ZZZ, $\mathrm{WW\gamma}$). These, along with $\mathrm{W^{\pm}Z}$ and ZZ, are considered to be irreducible. While most of these processes have extremely small cross-sections, they cannot be neglected when estimating the total background to this search. Predictions based on Madgraph-generated Monte Carlo samples are used to account for these rare processes. These MC samples do not include trigger inefficiencies and therefore a scale factor to take the trigger inefficiencies into account (described in Section 7.3) are applied. In addition, an event-by-event scale factor is applied to account for the b-tagging efficiency difference in data and MC. A difference in ID/isolation efficiencies was also observed between data and MC, so another scale factor is applied, described in Section 7.6.3.

Tables 7-15 to 7-18 show the irreducible background from each process for four baseline regions as examples.

Table 7-15. L0 irreducible background obtained with MC samples.

Samples	$\mu\mu$	$e\mu$	ee	Total
$t\bar{t}W$	10.71 $\pm$ 5.38	16.90 $\pm$ 8.47	6.68 $\pm$ 3.37	34.29 $\pm$ 17.17
WZ	7.50 $\pm$ 3.76	15.41 $\pm$ 7.71	6.59 $\pm$ 3.31	29.50 $\pm$ 14.76
$t\bar{t}Z$	2.73 $\pm$ 1.39	4.91 $\pm$ 2.48	1.90 $\pm$ 0.97	9.54 $\pm$ 4.79
WWW	1.43 $\pm$ 0.72	2.26 $\pm$ 1.14	1.04 $\pm$ 0.53	4.73 $\pm$ 2.37
$t\bar{t}\gamma$	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	3.58 $\pm$ 2.21
ZZ	0.38 $\pm$ 0.19	0.83 $\pm$ 0.41	0.42 $\pm$ 0.21	1.63 $\pm$ 0.81
WWZ	0.23 $\pm$ 0.12	0.45 $\pm$ 0.23	0.23 $\pm$ 0.12	0.91 $\pm$ 0.46
$t\bar{t}WW$	0.29 $\pm$ 0.14	0.45 $\pm$ 0.22	0.19 $\pm$ 0.09	0.93 $\pm$ 0.46
ZZZ	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.01 $\pm$ 0.00
WZZ	0.06 $\pm$ 0.03	0.14 $\pm$ 0.07	0.06 $\pm$ 0.03	0.26 $\pm$ 0.13
W^+W^+	6.65 $\pm$ 3.42	10.34 $\pm$ 5.26	3.28 $\pm$ 1.74	20.27 $\pm$ 10.22
W^-W^-	1.76 $\pm$ 0.90	3.09 $\pm$ 1.57	1.15 $\pm$ 0.60	6.00 $\pm$ 3.02
WW (DPS)	0.01 $\pm$ 0.03	0.02 $\pm$ 0.03	0.01 $\pm$ 0.03	0.05 $\pm$ 0.05
tbZ	0.18 $\pm$ 0.09	0.41 $\pm$ 0.21	0.20 $\pm$ 0.10	0.80 $\pm$ 0.40
$WW\gamma$	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09
$W\gamma^* \to \mu\mu$	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19
$W\gamma^* \to \tau\tau$	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24
$H \to WW$	2.39 $\pm$ 1.22	4.07 $\pm$ 2.06	1.84 $\pm$ 0.95	8.29 $\pm$ 4.17
$H \to ZZ$	0.12 $\pm$ 0.06	0.22 $\pm$ 0.11	0.09 $\pm$ 0.05	0.43 $\pm$ 0.22
$H \to \tau\tau$	0.22 $\pm$ 0.11	0.29 $\pm$ 0.15	0.15 $\pm$ 0.08	0.66 $\pm$ 0.33
Total	34.67 $\pm$ 17.37	59.79 $\pm$ 29.93	23.82 $\pm$ 11.94	121.86 $\pm$ 61.00

Table 7-16. L2 irreducible background obtained with MC samples.

Samples	$\mu\mu$	$e\mu$	ee	Total
$t\bar{t}W$	3.24 ± 1.65	5.33 ± 2.69	2.11 ± 1.08	10.68 ± 5.37
WZ	0.03 ± 0.03	0.05 ± 0.04	0.05 ± 0.04	0.13 ± 0.08
$t\bar{t}Z$	0.88 ± 0.47	1.55 ± 0.80	0.62 ± 0.33	3.06 ± 1.55
WWW	0.01 ± 0.02	0.01 ± 0.02	0.01 ± 0.02	0.04 ± 0.03
$t\bar{t}\gamma$	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	1.24 ± 0.77
ZZ	0.00 ± 0.00	0.01 ± 0.00	0.00 ± 0.00	0.01 ± 0.01
WWZ	0.01 ± 0.01	0.01 ± 0.01	0.00 ± 0.01	0.03 ± 0.02
$t\bar{t}WW$	0.09 ± 0.04	0.13 ± 0.07	0.06 ± 0.03	0.27 ± 0.14
ZZZ	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
WZZ	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.01 ± 0.01
W^+W^+	0.00 ± 0.14	0.06 ± 0.14	0.00 ± 0.14	0.06 ± 0.14
W^-W^-	0.02 ± 0.04	0.03 ± 0.05	0.00 ± 0.03	0.05 ± 0.06
WW (DPS)	0.00 ± 0.03	0.00 ± 0.03	0.00 ± 0.03	0.00 ± 0.03
tbZ	0.04 ± 0.02	0.07 ± 0.04	0.03 ± 0.02	0.13 ± 0.07
$WW\gamma$	0.00 ± 0.09	0.00 ± 0.09	0.00 ± 0.09	0.00 ± 0.09
$W\gamma^* \to \mu\mu$	0.00 ± 0.19	0.00 ± 0.19	0.00 ± 0.19	0.00 ± 0.19
$W\gamma^* \to \tau\tau$	0.00 ± 0.24	0.00 ± 0.24	0.00 ± 0.24	0.00 ± 0.24
$H \to WW$	0.66 ± 0.36	1.05 ± 0.56	0.58 ± 0.32	2.29 ± 1.17
$H \to ZZ$	0.04 ± 0.02	0.07 ± 0.04	0.02 ± 0.01	0.14 ± 0.07
$H \to \tau\tau$	0.06 ± 0.03	0.08 ± 0.04	0.02 ± 0.02	0.15 ± 0.08
Total	5.07 ± 2.59	8.47 ± 4.27	3.51 ± 1.82	18.29 ± 9.21

Table 7-17. H0 irreducible background obtained with MC samples.

Samples	$\mu\mu$	$e\mu$	ee	Total
$t\bar{t}W$	14.82 $\pm$ 7.44	26.65 $\pm$ 13.35	10.69 $\pm$ 5.37	52.16 $\pm$ 26.10
WZ	17.24 $\pm$ 8.63	37.06 $\pm$ 18.54	18.70 $\pm$ 9.36	73.00 $\pm$ 36.51
$t\bar{t}Z$	3.19 $\pm$ 1.62	6.51 $\pm$ 3.28	2.90 $\pm$ 1.47	12.59 $\pm$ 6.32
WWW	2.44 $\pm$ 1.23	4.22 $\pm$ 2.12	1.86 $\pm$ 0.94	8.52 $\pm$ 4.27
$t\bar{t}\gamma$	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	3.58 $\pm$ 2.21
ZZ	0.91 $\pm$ 0.46	2.25 $\pm$ 1.13	1.22 $\pm$ 0.61	4.38 $\pm$ 2.19
WWZ	0.32 $\pm$ 0.17	0.81 $\pm$ 0.41	0.40 $\pm$ 0.20	1.53 $\pm$ 0.77
$t\bar{t}WW$	0.29 $\pm$ 0.15	0.51 $\pm$ 0.26	0.22 $\pm$ 0.11	1.02 $\pm$ 0.51
ZZZ	0.00 $\pm$ 0.00	0.01 $\pm$ 0.00	0.01 $\pm$ 0.01	0.02 $\pm$ 0.01
WZZ	0.10 $\pm$ 0.05	0.22 $\pm$ 0.11	0.11 $\pm$ 0.06	0.42 $\pm$ 0.21
W^+W^+	8.78 $\pm$ 4.47	14.96 $\pm$ 7.56	5.78 $\pm$ 2.97	29.52 $\pm$ 14.83
W^-W^-	2.90 $\pm$ 1.47	5.43 $\pm$ 2.74	2.11 $\pm$ 1.08	10.44 $\pm$ 5.24
WW (DPS)	0.23 $\pm$ 0.14	0.33 $\pm$ 0.19	0.13 $\pm$ 0.08	0.69 $\pm$ 0.36
tbZ	0.34 $\pm$ 0.17	0.83 $\pm$ 0.42	0.44 $\pm$ 0.22	1.61 $\pm$ 0.81
WWγ	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09
$W\gamma^* \rightarrow \mu\mu$	0.26 $\pm$ 0.30	0.36 $\pm$ 0.35	0.00 $\pm$ 0.19	0.62 $\pm$ 0.46
$W\gamma^* \rightarrow \tau\tau$	0.22 $\pm$ 0.32	0.00 $\pm$ 0.24	0.11 $\pm$ 0.25	0.33 $\pm$ 0.38
$H \rightarrow WW$	2.97 $\pm$ 1.51	5.92 $\pm$ 2.99	2.70 $\pm$ 1.38	11.58 $\pm$ 5.82
$H \rightarrow ZZ$	0.15 $\pm$ 0.08	0.28 $\pm$ 0.14	0.12 $\pm$ 0.06	0.55 $\pm$ 0.28
$H \rightarrow \tau\tau$	0.43 $\pm$ 0.22	0.72 $\pm$ 0.36	0.38 $\pm$ 0.19	1.53 $\pm$ 0.77
Total	55.59 $\pm$ 27.82	107.06 $\pm$ 53.56	47.87 $\pm$ 23.96	214.10 $\pm$ 107.10

Table 7-18. H2 irreducible background obtained with MC samples.

Samples	$\mu\mu$	$e\mu$	ee	Total
$t\bar{t}W$	3.89 $\pm$ 1.97	7.30 $\pm$ 3.68	2.89 $\pm$ 1.47	14.08 $\pm$ 7.07
WZ	0.08 $\pm$ 0.06	0.14 $\pm$ 0.09	0.04 $\pm$ 0.04	0.27 $\pm$ 0.15
$t\bar{t}Z$	0.71 $\pm$ 0.38	1.83 $\pm$ 0.94	0.65 $\pm$ 0.35	3.19 $\pm$ 1.61
WWW	0.02 $\pm$ 0.02	0.04 $\pm$ 0.03	0.01 $\pm$ 0.01	0.07 $\pm$ 0.04
$t\bar{t}\gamma$	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.99 $\pm$ 0.61
ZZ	0.01 $\pm$ 0.00	0.02 $\pm$ 0.01	0.00 $\pm$ 0.00	0.03 $\pm$ 0.01
WWZ	0.01 $\pm$ 0.01	0.02 $\pm$ 0.02	0.00 $\pm$ 0.01	0.03 $\pm$ 0.02
$t\bar{t}WW$	0.07 $\pm$ 0.04	0.13 $\pm$ 0.07	0.06 $\pm$ 0.03	0.26 $\pm$ 0.13
ZZZ	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00	0.00 $\pm$ 0.00
WZZ	0.00 $\pm$ 0.01	0.01 $\pm$ 0.01	0.01 $\pm$ 0.01	0.02 $\pm$ 0.01
W^+W^+	0.00 $\pm$ 0.13	0.06 $\pm$ 0.13	0.00 $\pm$ 0.13	0.06 $\pm$ 0.13
W^-W^-	0.03 $\pm$ 0.04	0.03 $\pm$ 0.04	0.00 $\pm$ 0.03	0.06 $\pm$ 0.06
WW (DPS)	0.00 $\pm$ 0.03	0.00 $\pm$ 0.03	0.00 $\pm$ 0.03	0.00 $\pm$ 0.03
tbZ	0.05 $\pm$ 0.02	0.09 $\pm$ 0.05	0.05 $\pm$ 0.03	0.19 $\pm$ 0.10
$WW\gamma$	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09	0.00 $\pm$ 0.09
$W\gamma^* \to \mu\mu$	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19	0.00 $\pm$ 0.19
$W\gamma^* \to \tau\tau$	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24	0.00 $\pm$ 0.24
$H \to WW$	0.63 $\pm$ 0.35	0.99 $\pm$ 0.53	0.56 $\pm$ 0.31	2.18 $\pm$ 1.12
$H \to ZZ$	0.03 $\pm$ 0.02	0.06 $\pm$ 0.03	0.02 $\pm$ 0.01	0.11 $\pm$ 0.06
$H \to \tau\tau$	0.06 $\pm$ 0.04	0.10 $\pm$ 0.06	0.05 $\pm$ 0.03	0.21 $\pm$ 0.11
Total	5.59 $\pm$ 2.84	10.83 $\pm$ 5.45	4.34 $\pm$ 2.23	21.76 $\pm$ 10.92

7.5.4 Summary of Results

Tables 7-19 to 7-22 show the predicted backgrounds from each source for individual channels in the four baseline regions. Note that background from nonprompt leptons makes up 30-60% of the total prediction as does the rare MC (irreducible) background. The background from charge mis-id is only 1-2% of the total predicted background.

Figure 7-19 shows the predicted SM background from various sources in each channel for the baseline regions with $\mathrm{H_T} > 200$ GeV, $\not\mathrm{E}_\mathrm{T} > 50$ GeV and $\mathrm{N_{b-jets}} \geq 0$ or $\mathrm{N_{b-jets}} \geq 2$. This is the same information from Tables 7-19 to 7-22, only shown in graphical form. The largest fraction of the background belongs to top-pair production with one prompt lepton from W and one lepton from b-quark decays. The total uncertainty on the expected background is shown with the hatched band. For these four baseline regions where a relatively larger number of backgrounds are expected, very good agreement is obtained between predicted background and observed number of events, demonstrating that the background prediction methods perform well.

Figures 7-20-7-23 show kinematic distributions of the SM background from various sources for the high-$\mathrm{p_T}$ and low-$\mathrm{p_T}$ baselines. Predicted backgrounds are shown with a different color for each category (same color code as in Figure 7-19 applies). Please note that the background prediction methods, in particular the nonprompt lepton background prediction method, are not designed to predict the full shape of the kinematic distributions of the background. Nevertheless, the predicted distributions are in good agreement with the observed data.

Table 7-23 and Fig. 7-24 show good agreement between the predicted and observed event yields in the exclusive search regions of the low-$\mathrm{p_T}$ and high-$\mathrm{p_T}$ analyses. No significant excesses are observed. These results were combined with two other major analysis efforts for a single CMS publication. They were used to conduct a shape analysis, comparing the distribution of predicted and observed events in the 4D phase space binned in the observables $\mathrm{N_{jets}}$, $\mathrm{N_{b-jets}}$, $\mathrm{H_T}$, and $\not\mathrm{E}_\mathrm{T}$.

Table 7-19. L0 yields and background predictions.

	$\mu\mu$	ee	eμ	Total	(%) of Total
Nonprompt	48.20 $\pm$ 1.48 $\pm$ 14.46	49.16 $\pm$ 2.16 $\pm$ 24.58	94.13 $\pm$ 2.72 $\pm$ 31.97	191.49 $\pm$ 3.77 $\pm$ 60.29	60.29 %
Charge-flip	0	3.31 $\pm$ 0.04 $\pm$ 0.66	0.98 $\pm$ 0.02 $\pm$ 0.20	4.29 $\pm$ 0.05 $\pm$ 0.86	1.35 %
Rare MC	34.67 $\pm$ 1.12 $\pm$ 17.34	23.82 $\pm$ 0.91 $\pm$ 11.91	59.79 $\pm$ 1.39 $\pm$ 29.90	121.86 $\pm$ 2.91 $\pm$ 60.93	38.36 %
TOTAL	82.88 $\pm$ 1.86 $\pm$ 22.58	76.29 $\pm$ 2.34 $\pm$ 27.32	154.90 $\pm$ 3.05 $\pm$ 43.77	317.64 $\pm$ 4.76 $\pm$ 85.72	100 %
OBSERVED	110	75	164	349	

Table 7-20. L2 yields and background predictions.

	$\mu\mu$	ee	eμ	Total	(%) of Total
Nonprompt	8.23 $\pm$ 0.86 $\pm$ 2.47	2.42 $\pm$ 0.41 $\pm$ 1.21	11.07 $\pm$ 0.99 $\pm$ 2.93	21.72 $\pm$ 1.37 $\pm$ 5.66	53.51 %
Charge-flip	0	0.28 $\pm$ 0.01 $\pm$ 0.06	0.29 $\pm$ 0.01 $\pm$ 0.06	0.58 $\pm$ 0.02 $\pm$ 0.12	1.42 %
Rare MC	5.07 $\pm$ 0.51 $\pm$ 2.54	3.51 $\pm$ 0.47 $\pm$ 1.75	8.47 $\pm$ 0.58 $\pm$ 4.23	18.29 $\pm$ 1.06 $\pm$ 9.14	45.07 %
TOTAL	13.30 $\pm$ 1.00 $\pm$ 3.54	6.21 $\pm$ 0.63 $\pm$ 2.13	19.83 $\pm$ 1.15 $\pm$ 5.15	40.58 $\pm$ 1.74 $\pm$ 10.75	100 %
OBSERVED	14	7	24	45	

Table 7-21. H0 yields and background predictions.

	$\mu\mu$	ee	eμ	Total	(%) of Total
Nonprompt	27.65 $\pm$ 1.17 $\pm$ 8.30	49.96 $\pm$ 2.37 $\pm$ 24.98	83.96 $\pm$ 2.85 $\pm$ 31.89	161.56 $\pm$ 3.89 $\pm$ 58.14	54.30 %
Charge-flip	0	4.43 $\pm$ 0.05 $\pm$ 0.89	1.27 $\pm$ 0.02 $\pm$ 0.25	5.69 $\pm$ 0.06 $\pm$ 1.14	1.91 %
Rare MC	33.72 $\pm$ 1.08 $\pm$ 16.86	28.16 $\pm$ 0.94 $\pm$ 14.08	64.81 $\pm$ 1.38 $\pm$ 32.41	130.27 $\pm$ 2.90 $\pm$ 65.13	43.78 %
TOTAL	61.37 $\pm$ 1.59 $\pm$ 18.79	82.54 $\pm$ 2.55 $\pm$ 28.69	150.04 $\pm$ 3.17 $\pm$ 45.47	297.53 $\pm$ 4.85 $\pm$ 87.32	100 %
OBSERVED	67	72	117	256	

Table 7-22. H2 yields and background predictions.

	$\mu\mu$	ee	eμ	Total	(%) of Total
Nonprompt	2.78 $\pm$ 0.51 $\pm$ 0.83	1.00 $\pm$ 0.31 $\pm$ 0.50	5.45 $\pm$ 0.72 $\pm$ 1.51	9.22 $\pm$ 0.93 $\pm$ 2.44	33.06 %
Charge-flip	0	0.36 $\pm$ 0.01 $\pm$ 0.07	0.38 $\pm$ 0.01 $\pm$ 0.08	0.74 $\pm$ 0.02 $\pm$ 0.15	2.66 %
Rare MC	4.49 $\pm$ 0.49 $\pm$ 2.24	3.74 $\pm$ 0.47 $\pm$ 1.87	8.72 $\pm$ 0.58 $\pm$ 4.36	17.93 $\pm$ 0.95 $\pm$ 8.97	64.28 %
TOTAL	7.26 $\pm$ 0.71 $\pm$ 2.39	5.09 $\pm$ 0.57 $\pm$ 1.94	14.55 $\pm$ 0.92 $\pm$ 4.61	27.90 $\pm$ 1.33 $\pm$ 9.29	100 %
OBSERVED	14	4	19	37	

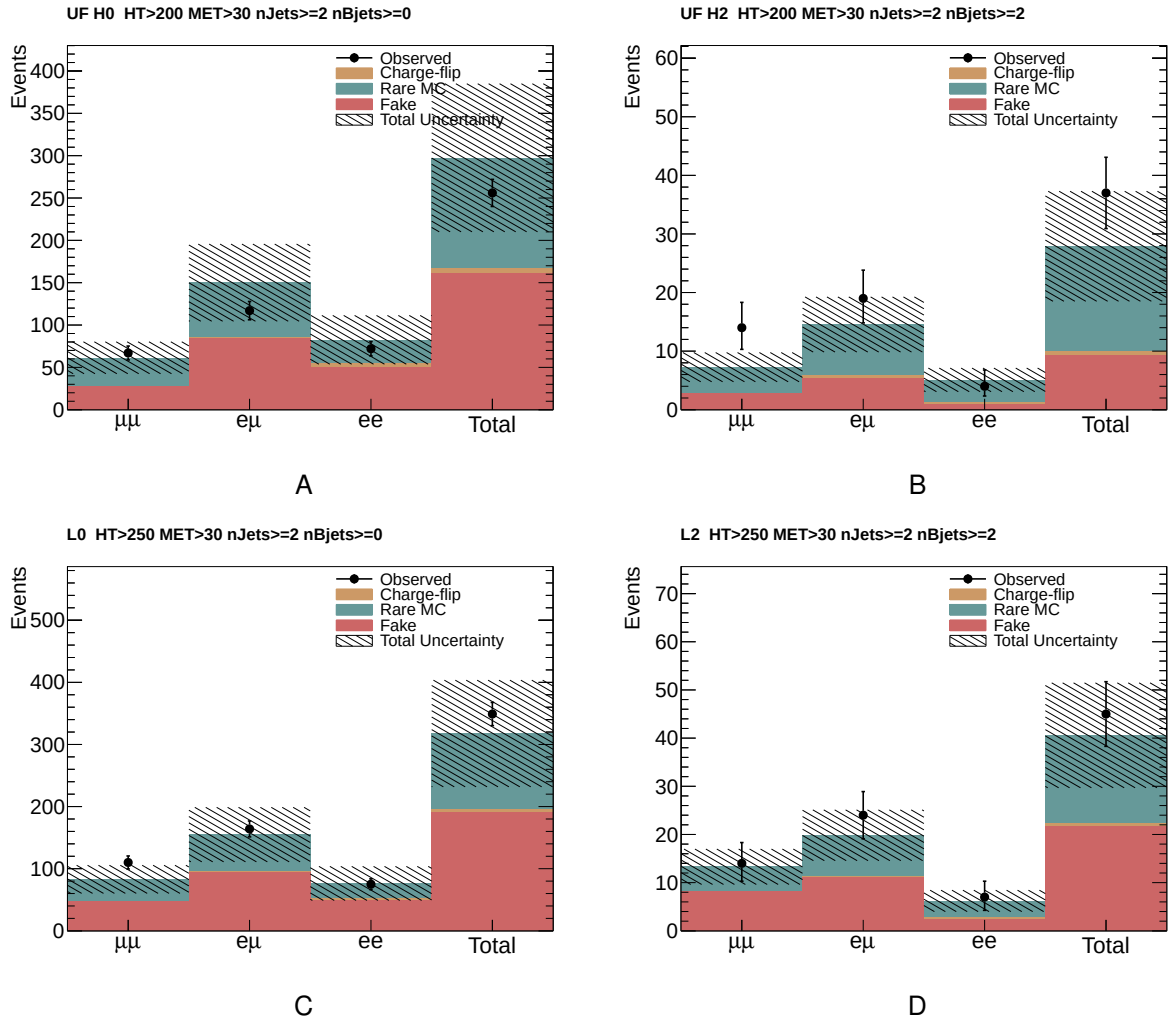


Figure 7-19. Baseline yields and background predictions in each channel. The results correspond to the high-$\mathrm{p_T}$ (A-B) and low-$\mathrm{p_T}$ (C-D) baseline selections where $\mathrm{N_{b-jets}} \geq 0$ (A, C) or $\mathrm{N_{b-jets}} \geq 2$ (B, D).

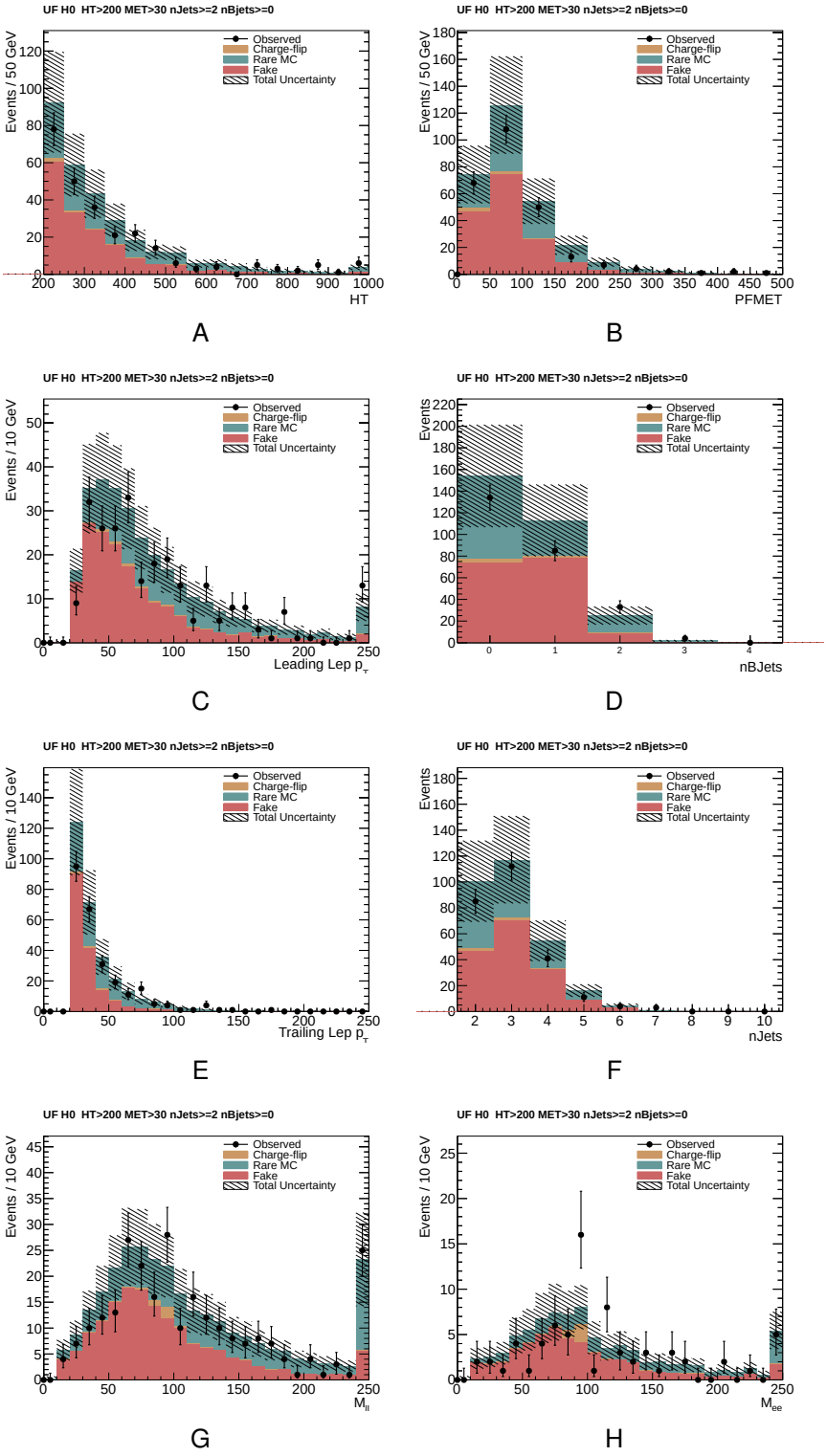


Figure 7-20. H0 kinematic distributions of predicted SM backgrounds and observed number of events are shown for: (A) $\mathrm{H_T}$, (B) $\not\!\!\mathrm{E_T}$, (C) leading lepton $\mathrm{p_T}$, (D) $\mathrm{N_{b-jets}}$, (E) trailing lepton $\mathrm{p_T}$, (F) $\mathrm{N_{jets}}$, (G) $\mathrm{M_{ll}}$, and (H) $\mathrm{M_{ee}}$ are plotted. The results correspond to the high-$\mathrm{p_T}$ baseline selection where $\mathrm{H_T} > 200$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV, $\mathrm{N_{b-jets}} \geq 0$.

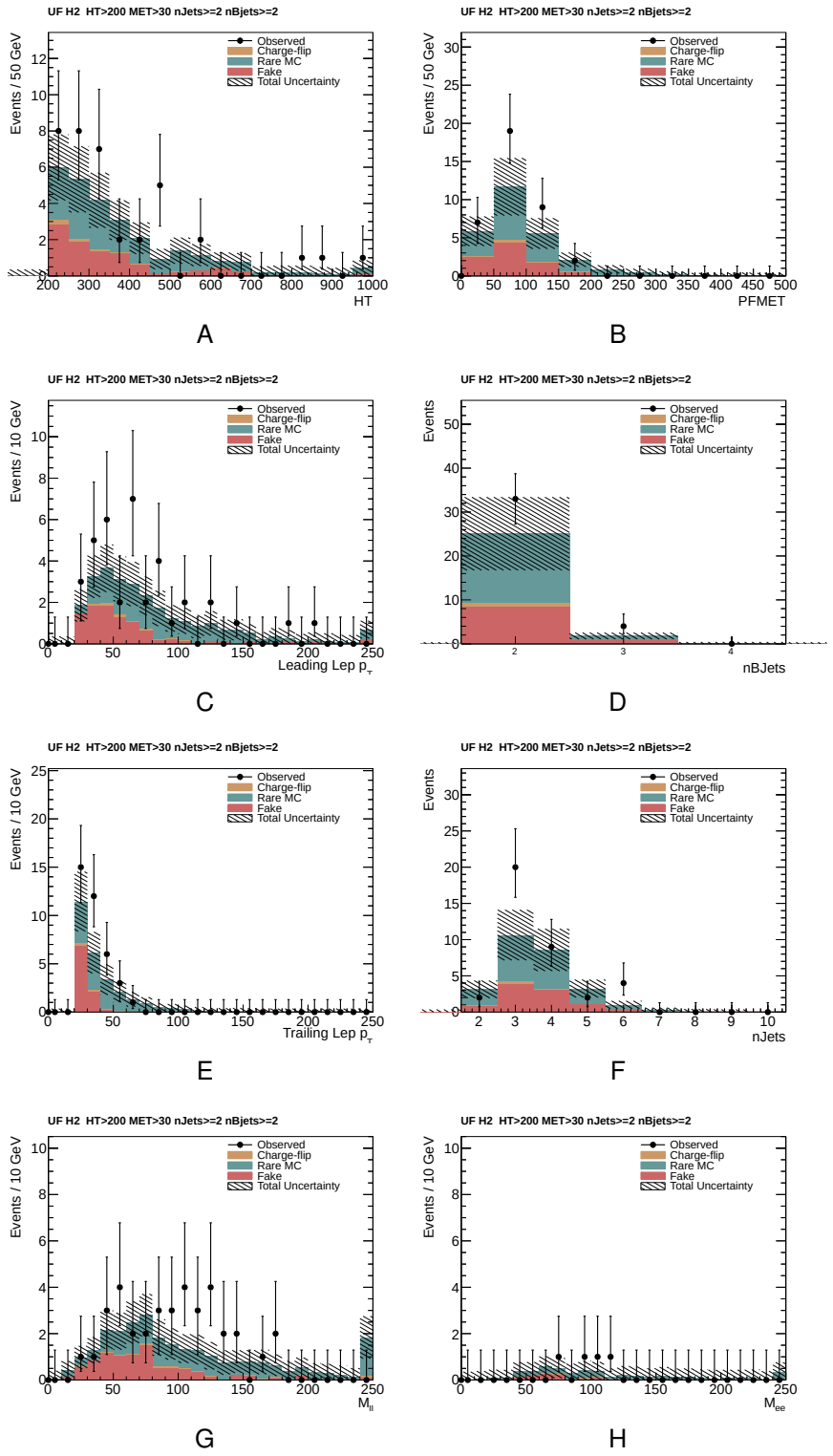


Figure 7-21. H2 kinematic distributions of predicted SM backgrounds and observed number of events are shown for: (A) $\mathrm{H_T}$, (B) $\not\!\!\mathrm{E_T}$, (C) leading lepton $\mathrm{p_T}$, (D) $\mathrm{N_{b-jets}}$, (E) trailing lepton $\mathrm{p_T}$, (F) $\mathrm{N_{jets}}$, (G) $\mathrm{M_{ll}}$, and (H) $\mathrm{M_{ee}}$ are plotted. The results correspond to the high-$\mathrm{p_T}$ baseline selection where $\mathrm{H_T} > 200$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV, $\mathrm{N_{b-jets}} \geq 2$.

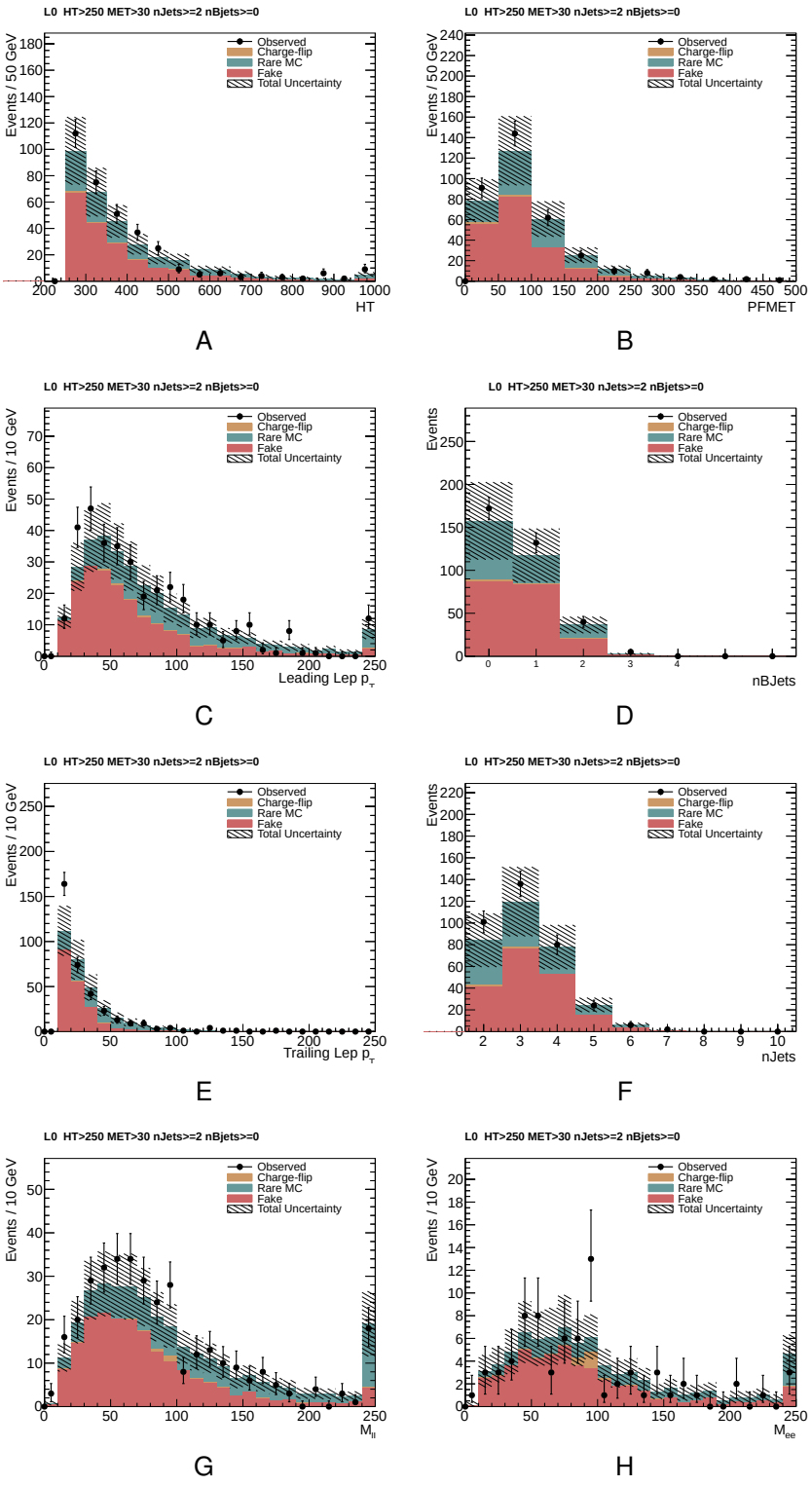


Figure 7-22. L0 kinematic distributions of predicted SM backgrounds and observed number of events are shown for: (A) $\mathrm{H_T}$, (B) $\not\!\!\mathrm{E_T}$, (C) leading lepton $\mathrm{p_T}$, (D) $\mathrm{N_{b-jets}}$, (E) trailing lepton $\mathrm{p_T}$, (F) $\mathrm{N_{jets}}$, (G) $\mathrm{M_{ll}}$, and (H) $\mathrm{M_{ee}}$ are plotted. The results correspond to the low-$\mathrm{p_T}$ baseline selection where $\mathrm{H_T} > 250$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV, $\mathrm{N_{b-jets}} \geq 0$.

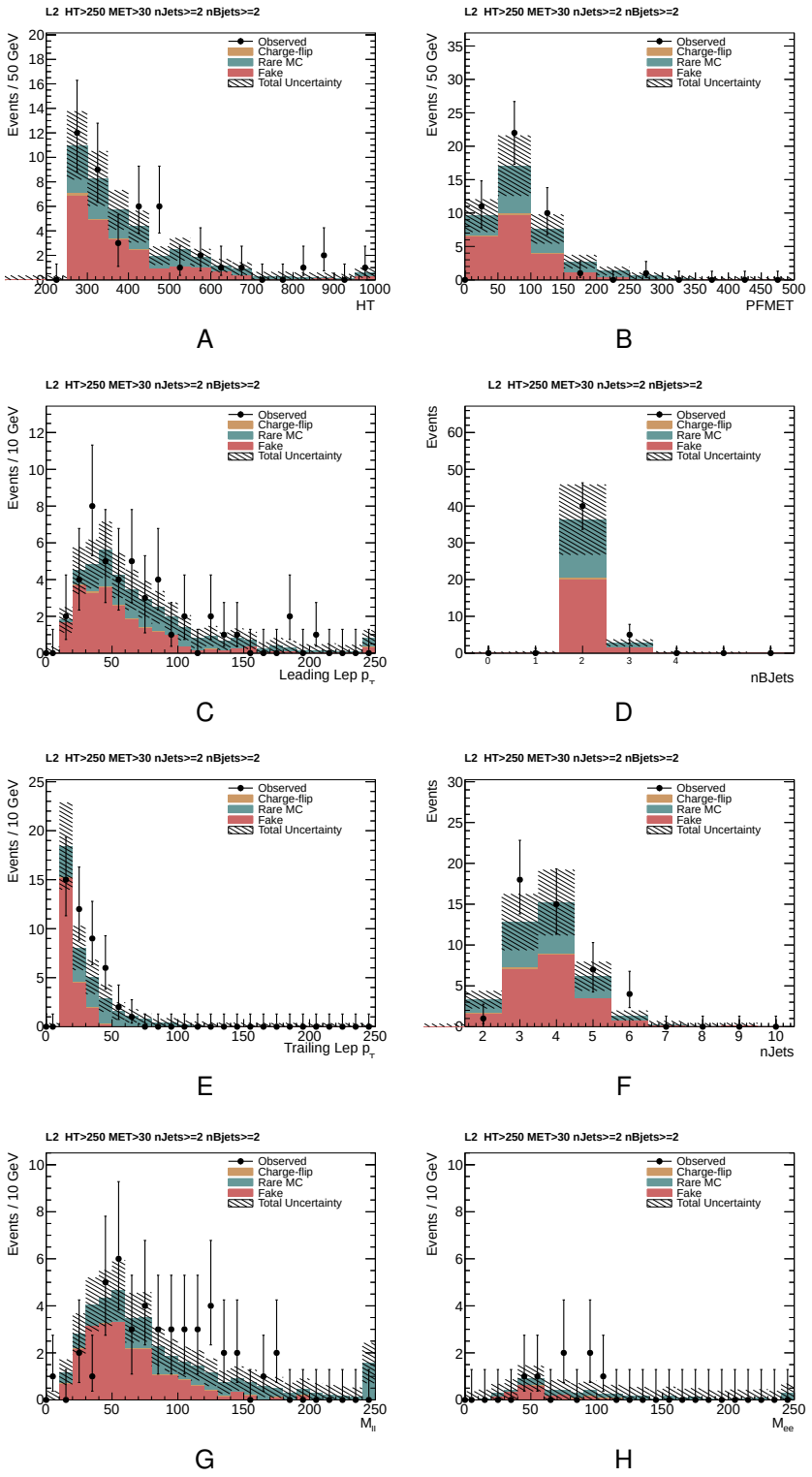


Figure 7-23. L2 kinematic distributions of predicted SM backgrounds and observed number of events are shown for: (A) $\mathrm{H_T}$, (B) $\not\!\!\mathrm{E_T}$, (C) leading lepton $\mathrm{p_T}$, (D) $\mathrm{N_{b-jets}}$, (E) trailing lepton $\mathrm{p_T}$, (F) $\mathrm{N_{jets}}$, (G) $\mathrm{M_{ll}}$, and (H) $\mathrm{M_{ee}}$ are plotted. The results correspond to the low-$\mathrm{p_T}$ baseline selection where $\mathrm{H_T} > 250$ GeV, $\not\!\!\mathrm{E_T} > 30$ GeV, $\mathrm{N_{b-jets}} \geq 2$.

Table 7-23. Low-$\mathrm{p_T}$ and high-$\mathrm{p_T}$ search yields and background predictions.

Low-$\mathrm{p_T}$ SR	Predicted	Observed	High-$\mathrm{p_T}$ SR	Predicted	Observed
1	44.71 $\pm$ 13.13	50	1	56.38 $\pm$ 17.52	48
2	12.21 $\pm$ 4.20	17	2	9.17 $\pm$ 3.50	11
3	13.07 $\pm$ 3.63	13	3	10.17 $\pm$ 3.00	5
4	10.35 $\pm$ 3.11	4	4	6.97 $\pm$ 2.30	2
5	22.13 $\pm$ 7.15	22	5	22.20 $\pm$ 7.63	12
6	12.55 $\pm$ 4.85	18	6	9.29 $\pm$ 3.88	11
7	3.92 $\pm$ 1.30	2	7	2.94 $\pm$ 1.07	1
8	6.25 $\pm$ 2.18	4	8	4.02 $\pm$ 1.57	3
11	30.45 $\pm$ 8.44	40	11	40.67 $\pm$ 12.60	29
12	5.85 $\pm$ 1.70	5	12	3.70 $\pm$ 1.23	5
13	15.65 $\pm$ 4.40	15	13	12.00 $\pm$ 3.56	6
14	9.98 $\pm$ 2.91	6	14	6.51 $\pm$ 2.08	2
15	12.52 $\pm$ 3.43	9	15	11.19 $\pm$ 3.31	11
16	5.30 $\pm$ 1.70	5	16	3.90 $\pm$ 1.40	2
17	3.94 $\pm$ 1.20	3	17	2.75 $\pm$ 0.95	3
18	6.77 $\pm$ 2.10	11	18	4.65 $\pm$ 1.60	7
21	7.73 $\pm$ 2.11	10	21	6.95 $\pm$ 2.27	12
22	1.52 $\pm$ 0.65	1	22	0.96 $\pm$ 0.54	1
23	7.37 $\pm$ 2.01	6	23	3.74 $\pm$ 1.30	3
24	4.65 $\pm$ 1.54	11	24	2.77 $\pm$ 1.18	7
25	2.83 $\pm$ 0.99	1	25	2.91 $\pm$ 1.10	4
26	1.27 $\pm$ 0.63	2	26	0.82 $\pm$ 0.54	1
27	1.87 $\pm$ 0.71	0	27	1.30 $\pm$ 0.61	0
28	3.38 $\pm$ 1.23	3	28	2.11 $\pm$ 0.95	2

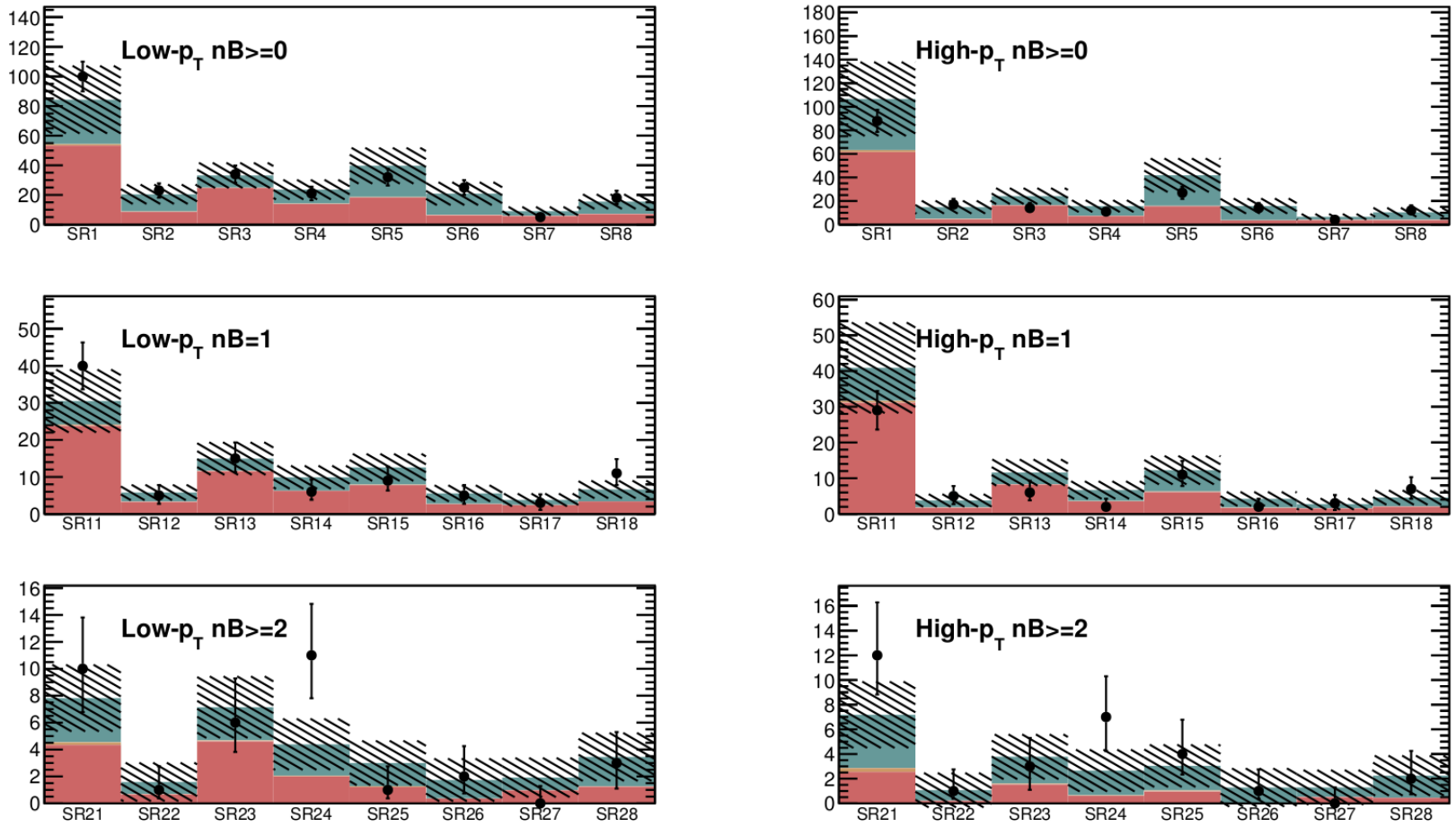


Figure 7-24. Plots of low-$\mathrm{p_T}$ and high-$\mathrm{p_T}$ search yields and background predictions are shown.

7.6 Signal Uncertainties

7.6.1 Theoretical Uncertainties

Apart from uncertainties particular to a given model of interest (e.g. cross-sections, branching ratios, etc), the main contributions to theoretical uncertainties are associated with the scale of the QCD coupling and the proton's parton density function.

For gluino/squark production in simplified SUSY models, the pdf uncertainties yield about a 2% systematic error on the $\mathrm{H_T}$ and $\not\mathrm{E}_\mathrm{T}$ acceptance [27], with some small dependence on the mass scales involved. The effect on the cross-section is expected to be much larger and has a strong dependence on the masses. These uncertainties were calculated centrally, as they were common to all analyses probing the same SUSY models for the CMS publication.

The systematic errors arising from the QCD scale are typically evaluated on CMS by varying it by a factor of two up and down, and observing the resulting change in cross-sections. This can be a model-dependent uncertainty, and thus was also evaluated centrally for the theoretical models used to interpret the results of this analysis.

7.6.2 Luminosity Uncertainty

The recommended uncertainty on the integrated luminosity for 2012 data was 4.5% [28].

7.6.3 Lepton Reconstruction, Identification, and Isolation Efficiencies

A full and thorough study of the lepton reconstruction, identification, and isolation modeling was performed using well-established tag-and-probe techniques on Z-control samples in data. The detailed description of this study as well as the results are available elsewhere [29]. The applied scale factors are summarized in Tables 7-24 and 7-25.

The systematic uncertainty (per lepton) to account for the modeling of the lepton acceptance is evaluated from the stability of the tag-and-probe results (Table 7-26). An

Table 7-24. Electron ID and isolation scale factors to be applied on the electrons in MC samples.

	p_T, GeV	$10-15$	$15-20$	$20-30$	$30-40$	$40-50$	$50-200$
η	$0.00-0.80$	0.834	0.918	0.954	0.960	0.972	0.969
	$0.80-1.44$	0.973	0.906	0.923	0.935	0.955	0.956
	$1.57-2.00$	0.954	0.909	0.921	0.924	0.950	0.995
	$2.00-2.50$	1.119	0.944	0.993	0.959	0.968	0.969

Table 7-25. Muon ID and isolation scale factors to be applied on the muons in MC samples.

	p_T, GeV	$10-15$	$15-20$	$20-30$	$30-40$	$40-50$	$50-200$
η	$0.00-1.20$	0.956	0.957	0.964	0.971	0.978	0.974
	$1.20-2.50$	0.960	0.971	0.981	0.978	0.984	0.977

additional source of uncertainty (per lepton) accounts for the fact that the modeling is assessed in Z-events (low jet activity) and then extrapolated to an environment with large jet activity (Table 7-27).

Table 7-26. Lepton scale factor uncertainty from tag & probe.

	$p_T < 15$ GeV	$p_T > 15$ GeV
e	10%	5%
μ	5%	3%

Table 7-27. Lepton scale factor uncertainty from event composition.

	$p_T < 30$ GeV	$p_T > 30$ GeV
e	3%	3%
μ	5%	3%

7.6.4 B-Jet Identification Efficiency

The probability was measured for a b-parton to be tagged with $\mathrm{CombinedSecondaryVertex}$ discriminator with the Medium Working point. The efficiency and its uncertainty have been conducted in [21, 22] with different methods on 2012 data. The efficiency is around 70% with systematic uncertainty at the level of 4%. Following the procedure described elsewhere [21], event-by-event scale factors were applied to the MC yields and the related uncertainties were extracted.

7.6.5 Hadronic Activity and MET Selection Efficiencies

The leading cause of uncertainty with respect to the $\mathrm{H_T}$ and $\not\mathrm{E}_\mathrm{T}$ efficiencies is due to the calibration of the jet energy scale (JES). CMS has demonstrated that calibrated Particle Flow jets with transverse momenta in the range of 40 GeV can have energy scale-uncertainties near 2-3% in the barrel and endcaps [30]. Therefore, a uniform uncertainty of 2.5% is assigned to the $\mathrm{p_T}$ measurements of all jets.

In order to evaluate how this JES uncertainty influences the signal acceptance, it is necessary to re-run the counting experiments in simulation after uniformly scaling up (and down) the candidate jet energies by 2.5%, and then propagating the effect to the $\mathrm{N_{jets}}$, $\mathrm{H_T}$, and $\not\mathrm{E}_\mathrm{T}$ observables. This is a model-dependent systematic uncertainty in general. Models with intrinsically high $\mathrm{H_T}$ scales (well above the respective $\mathrm{H_T}$ requirements) will hardly be influenced by fluctuations in the jet energy scale. However, models with characteristic $\mathrm{H_T}$ scales near the $\mathrm{H_T}$ requirements can be impacted significantly. This is an important consideration to remember when setting upper-limits. Figure 7-25 shows the results from fluctuating the JES up and down by 2.5% uniformly for all jets. The resulting uncertainty on the signal acceptance ranges from a few percent to 10%, depending on the mass splittings.

7.6.6 Trigger Uncertainties

Given the ample statistics captured with the control triggers, the efficiencies of the signal triggers were measured with a precision of $\pm$6%. To test models against the findings of the SS analysis, signal yields should be uniformly scaled down for each channel by the measured trigger efficiencies, which are reported in Section 7.3.

7.6.7 Summary of Signal Acceptance Uncertainties

A summary of the systematic uncertainties associated with the signal acceptance for this analysis is provided in Table 7-28. While a few uncertainties cannot be quantified until a signal model of interest is specified and the final state topologies and kinematics are studied, a lower bound on the signal acceptance uncertainty of 16.4% is given.

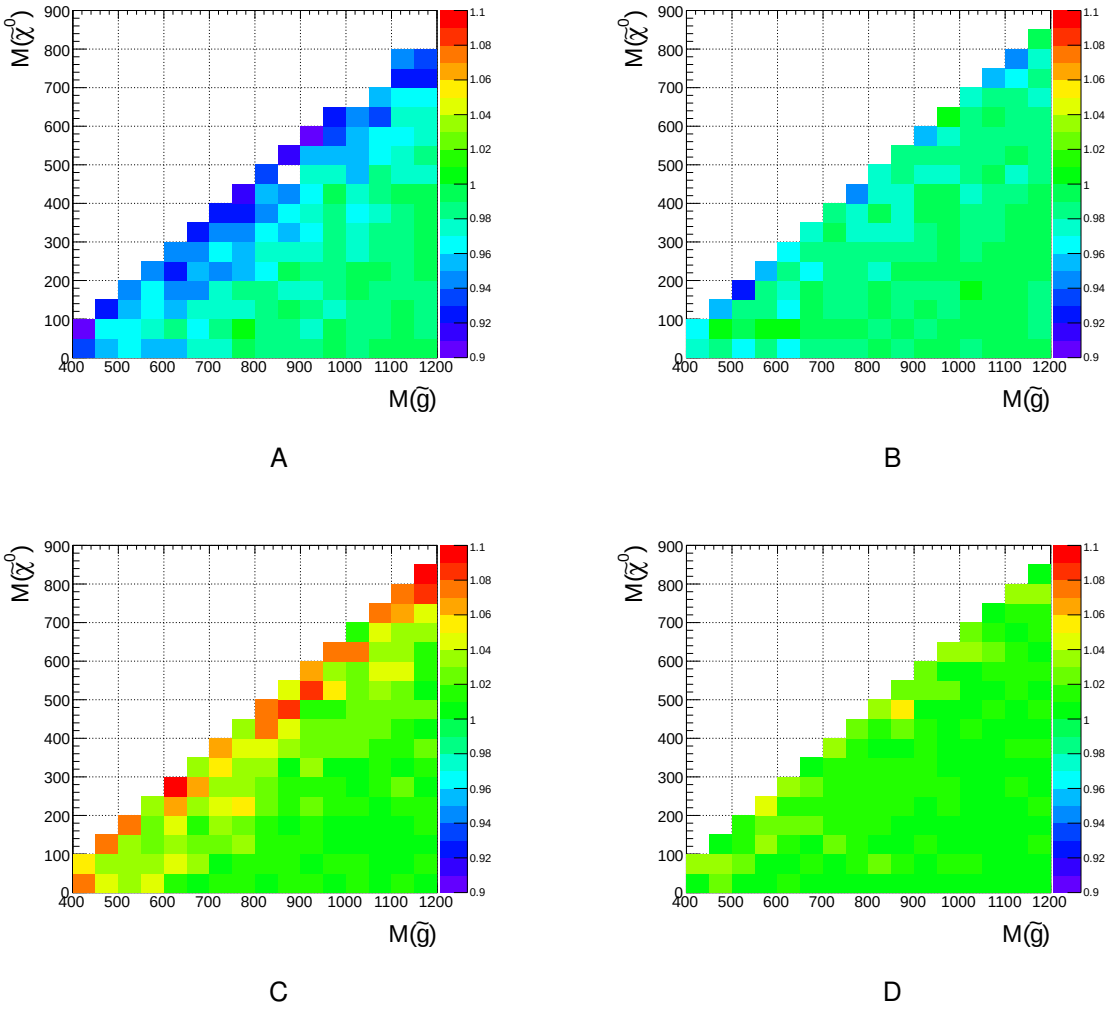


Figure 7-25. Effect of the JES fluctuation on signal acceptance of the SMS-T1tttt model for $H_T > 300$ GeV and $\not{E}_T > 50$ GeV (A, C) and $\not{E}_T > 120$ GeV (B, D). The uncertainties range from a few percent across the vast majority of the parameter space and rises to about 10% near the diagonals.

Table 7-28. Systematic uncertainties for the signal acceptance.

Source	%
Luminosity	4.5%
QCD scale	model dependent
PDF uncertainties	($\geq 2\%$)
Modeling of lepton reco, Id, iso based on Z-events	6%
Modeling SUSY lepton acceptance w/ Z-events	10%
Jet energy scale	10%
b-jet identification	4%
Trigger scaling	6%
Min. Systematic Error	16.4%

7.6.8 Signal Acceptance Model

To make the results of this analysis applicable in the context of other models, efficiencies are presented for lepton, missing energy , $\mathrm{H_T}$ and b-jet selection. In Fig. 7-26A lepton reconstruction and identification efficiencies are presented. In Fig. 7-26B, b-jet selection efficiencies are shown for the jets originating from different types of partons as a function of parton $\mathrm{p_T}$. While Fig. 7-26C and Fig. 7-26D show the $\not\mathrm{E}_\mathrm{T}$ and $\mathrm{H_T}$ selection efficiencies respectively. The validation of these efficiencies are performed using the MC simulated T1tttt model and presented in Fig 7-27.

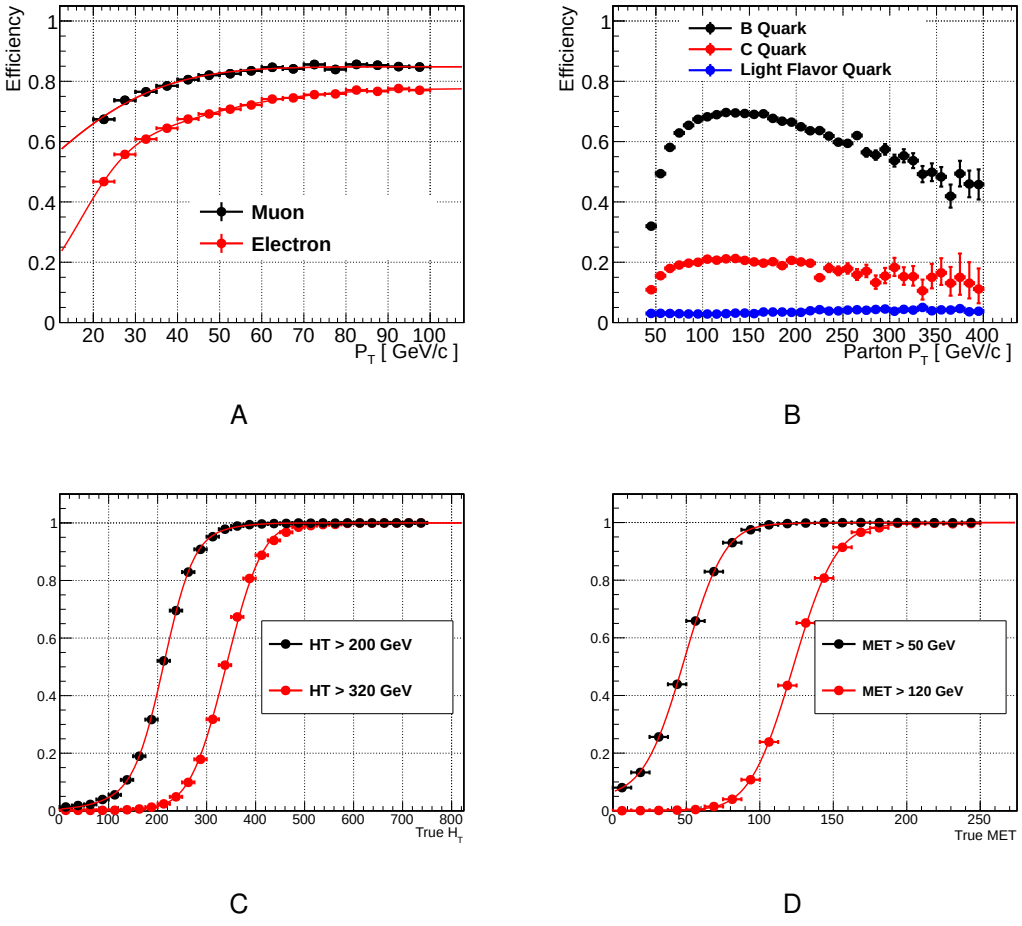


Figure 7-26. Signal efficiency model for (A) lepton, (B) b-jet, (C) $\mathrm{H_T}$, and (D) $\not\mathrm{E}_\mathrm{T}$ where selections are obtained in the $\mathrm{t\bar{t}}$ MC sample.

Those who are outside the CMS collaboration lack access to proprietary software for simulating the detector acceptance and the analysis acceptance. One can instead use MC-generated events in various SUSY models and apply the plots in Fig. 7-26 to find out whether that model is excluded. The closure test of this acceptance model is shown in Fig. 7-27.

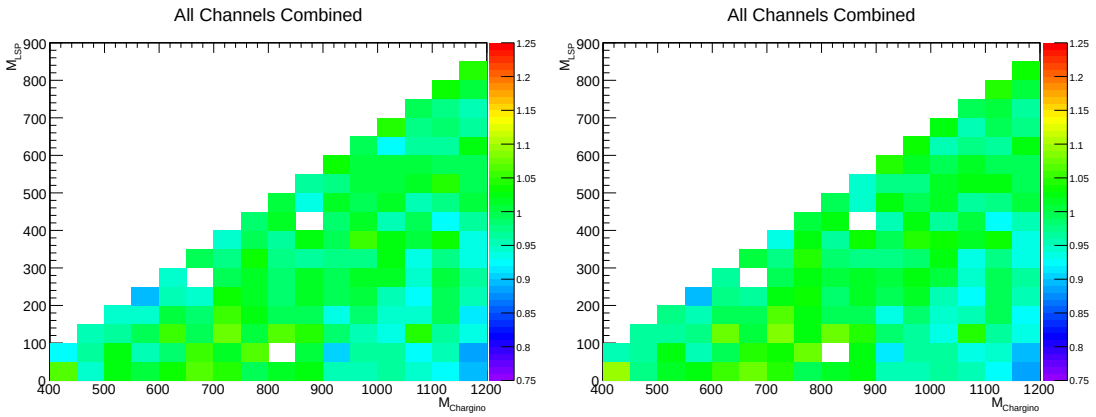


Figure 7-27. Closure test of the acceptance model parameterization. The ratio of the event yields per grid point obtained using the acceptance parameterization and the CMS detector simulation for the T1tttt model. The plot on the left shows this ratio for only the dilepton selection, while the plot on the right shows the closure test for the combined dilepton, $\mathrm{H_T} > 200$ GeV and $\not\!\!\mathrm{E_T} > 50$ GeV selection.

7.7 Interpretation of Results

The results of this thesis were combined with results from two other research groups within the CMS collaboration, who also analyzed the same-sign dilepton channel. Data-driven background estimation methods differed between the three groups, while signal event selection was the same. Good agreement between all three groups lends considerable confidence in the final results. Having observed no significant excesses above the Standard Model predictions in the combined result, a calculation of upper limits at 95% C.L. is performed on the simplified SUSY models introduced in Figs. 3-4-3.4. Model parameters are defined as in Table 7-29. In addition, same-sign dilepton analysis results were used to set limits on RPV models and same-sign top

Table 7-29. Simplified SUSY models and their parameters.

Signal Model	$m_{\tilde{g}}$	$m_{\tilde{t}}$	$m_{\tilde{b}}$	$m_{\tilde{\chi}^{\pm}}$	$m_{\tilde{\chi}^{0}}$
Model A1	[400, 1400]	inf	–	–	[1, 1200]
Model A2	[800, 1200]	[225, $m_{\tilde{g}}-175$]	–	–	50
Model B1	–	–	[325, 700]	[150, $m_{\tilde{b}}-175$]	50
	–	–	[200, 700]	[50, $m_{\tilde{b}}-100$]	[25, $0.5\times m_{\tilde{b}}-50$]
	–	–	[200, 700]	[32, $m_{\tilde{b}}-100$]	[25, $0.8\times m_{\tilde{b}}-80$]
Model B2	[800, 1400]	–	[400, 1400]	150	50
	[800, 1400]	–	[400, 1400]	300	50
Model C1	[800,1325]	–	–	[160, $m_{\tilde{g}}-80$]	[1, $m_{\tilde{g}}-100$]
	[400,1400]	–	–	[320, 1360]	[1, 1200]

production, $\sigma(\mathrm{pp} \to \mathrm{tt})$ and $\sigma(\mathrm{pp} \to \mathrm{tttt})$, though the UF analysis was not applicable and those results are not shown here.

The excluded phase space of models A1, A2, B1, B2, and C1 are shown in Figs. 7-28-7-30. Each grid point is treated as an individual counting experiment and excluded by the search region which sets the best limit. The limit is calculated with the LandS software [31] using the modified frequentist CL_s method. The same-sign dilepton analysis was designed to search for strongly-produced SUSY events, thus limits are shown in terms of $\tilde{\mathrm{g}}$ and $\tilde{\mathrm{b}}$ mass. The $\tilde{\mathrm{g}}$ mass is probed up to about 1050 GeV, as shown in Fig. 7-28, while the $\tilde{\mathrm{b}}$ mass is probed up to about 550 GeV, seen in Fig. 7-29. Even in the light squark scenario of model C1, $\tilde{\mathrm{g}}$ is excluded up to around 900 GeV.

Within the CMS collaboration, the SS dilepton analysis of the full 2012 dataset was denoted SUS-13-013. Figure 7.7 shows the $\tilde{\mathrm{g}}$ and $\tilde{\mathrm{b}}$ mass limits set by the SS analysis compared to other analyses, where SS dileptons is given (A) in green, and (B) in grey [32].

SUSY analyses within the CMS collaboration were kept as general as possible to allow for new physics discoveries, arising from SUSY or not. Across all the SUSY searches, no unambiguous hints of new physics have been observed. Thus, limits on many SUSY particle masses were presented by the CMS collaboration on simplified

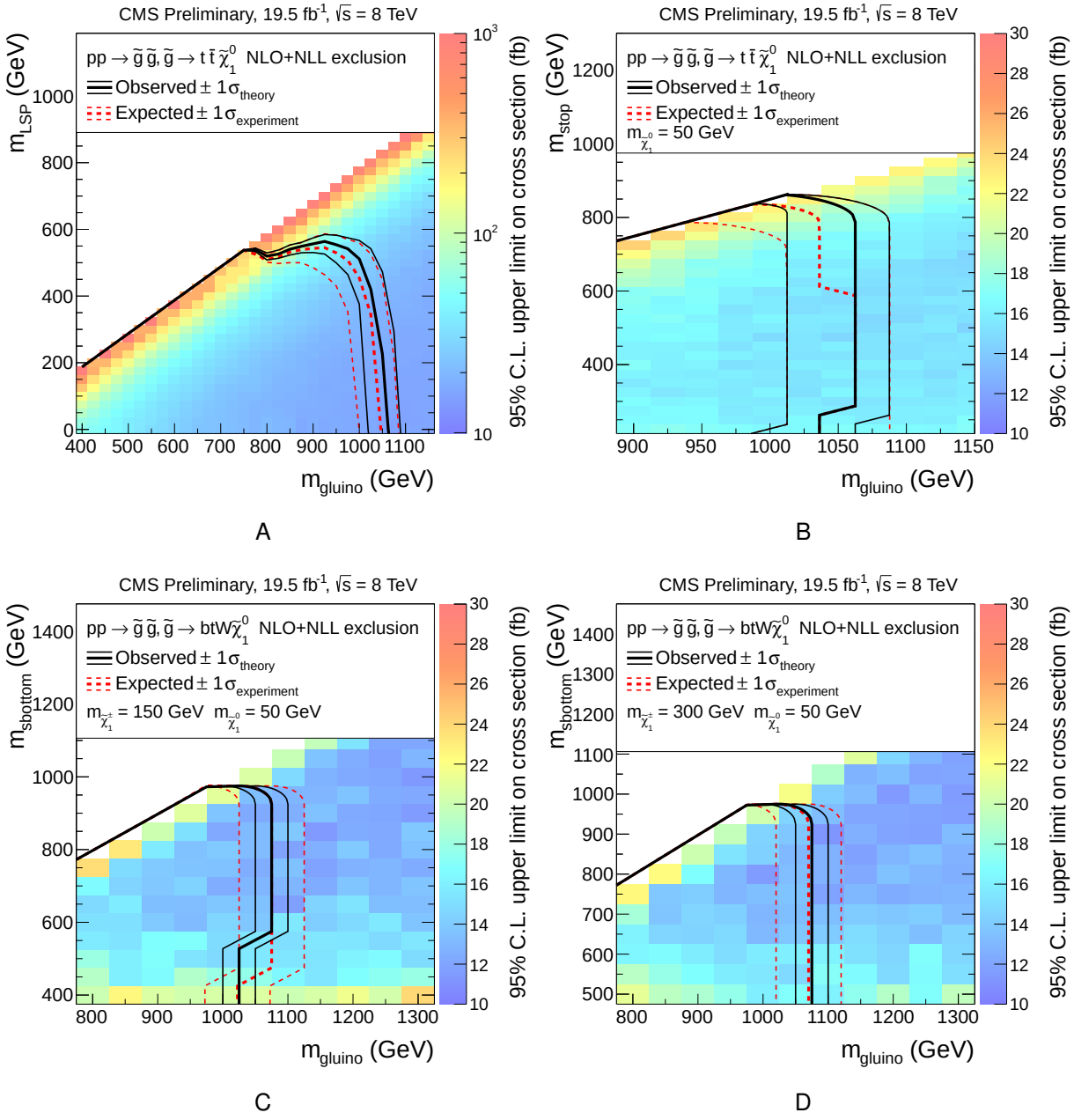


Figure 7-28. Excluded pointed for A and B models with direct gluino production: (A) model A1, (B) model A2 with 50 GeV LSP, (C) model B2 with 50 GeV LSP and 150 GeV chargino, and (D) model B2 with 50 GeV LSP and 300 GeV chargino.

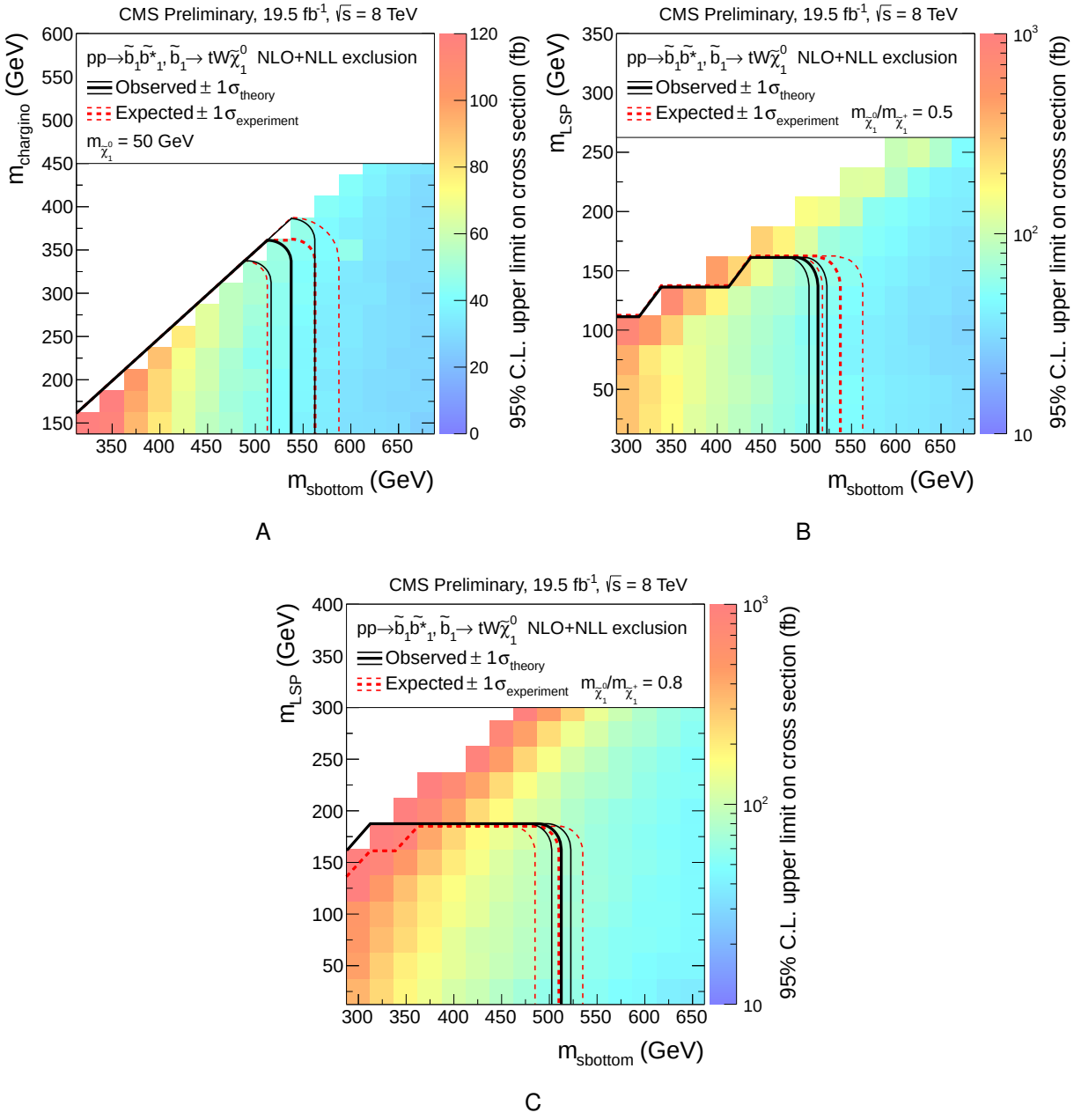


Figure 7-29. Excluded points for B models with direct sbottom production: (A) model B1 with 50 GeV LSP, (B) model B1 with m/m $= 0.5$, and (C) model B1 with m/m $= 0.8$.

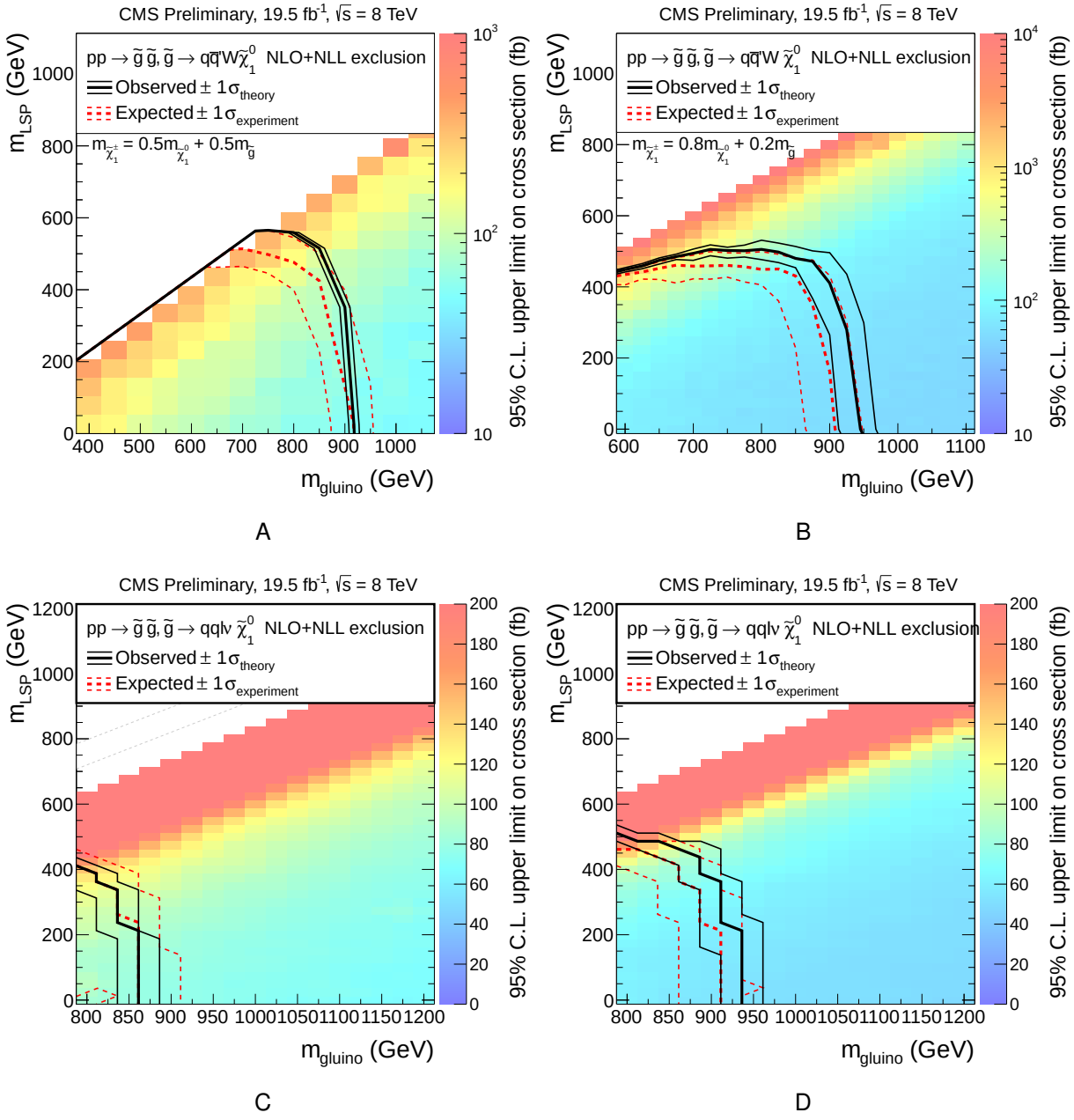
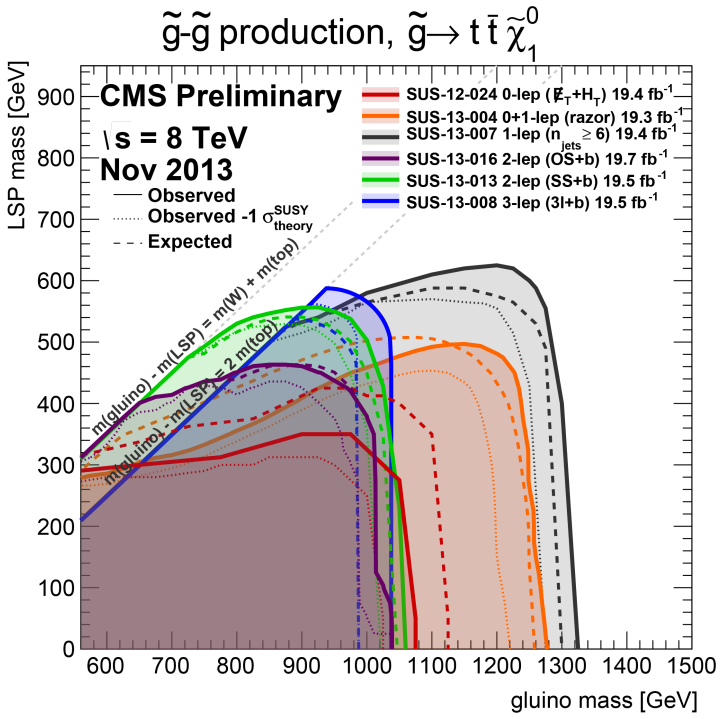
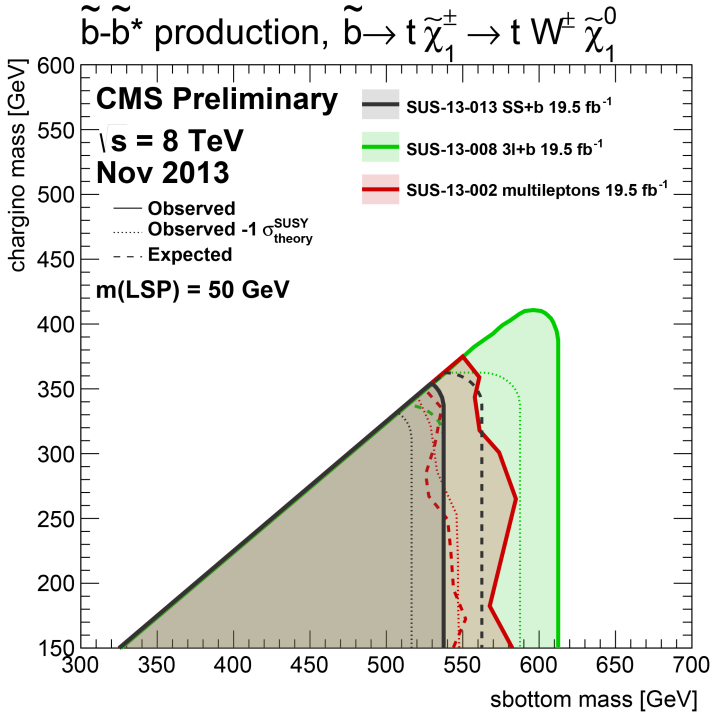


Figure 7-30. Excluded points for C models with direct gluino production: (A) with x=0.5: high-p_T signal regions with b-veto; (B) with x=0.8: low-p_T signal regions with b-veto; (C) with x=0.8: low-p_T signal regions with b-tags; (D) with x=0.8: low-p_T signal regions with b-veto.



A



B

Figure 7-31. Limits on sparticle masses set by the same-sign dilepton analysis, i.e. SUS-13-013, among others. Limits are shown for simplified SUSY models in green for (A) gluino mass and grey for (B) sbottom mass.

SUSY models. In these models, SUSY mass parameters are set to reasonable values so limits can be presented in at most a two-dimensional phase space. Thus, these limits only apply within the many assumptions of the simplified models. Fig. 7-32 shows selected limits on the masses of several sparticle categories where R-parity conservation is assumed [32]. The analysis of this thesis sets limits on gluino and sbottom mass, and these results are shown together with the limits set by other published CMS analyses.

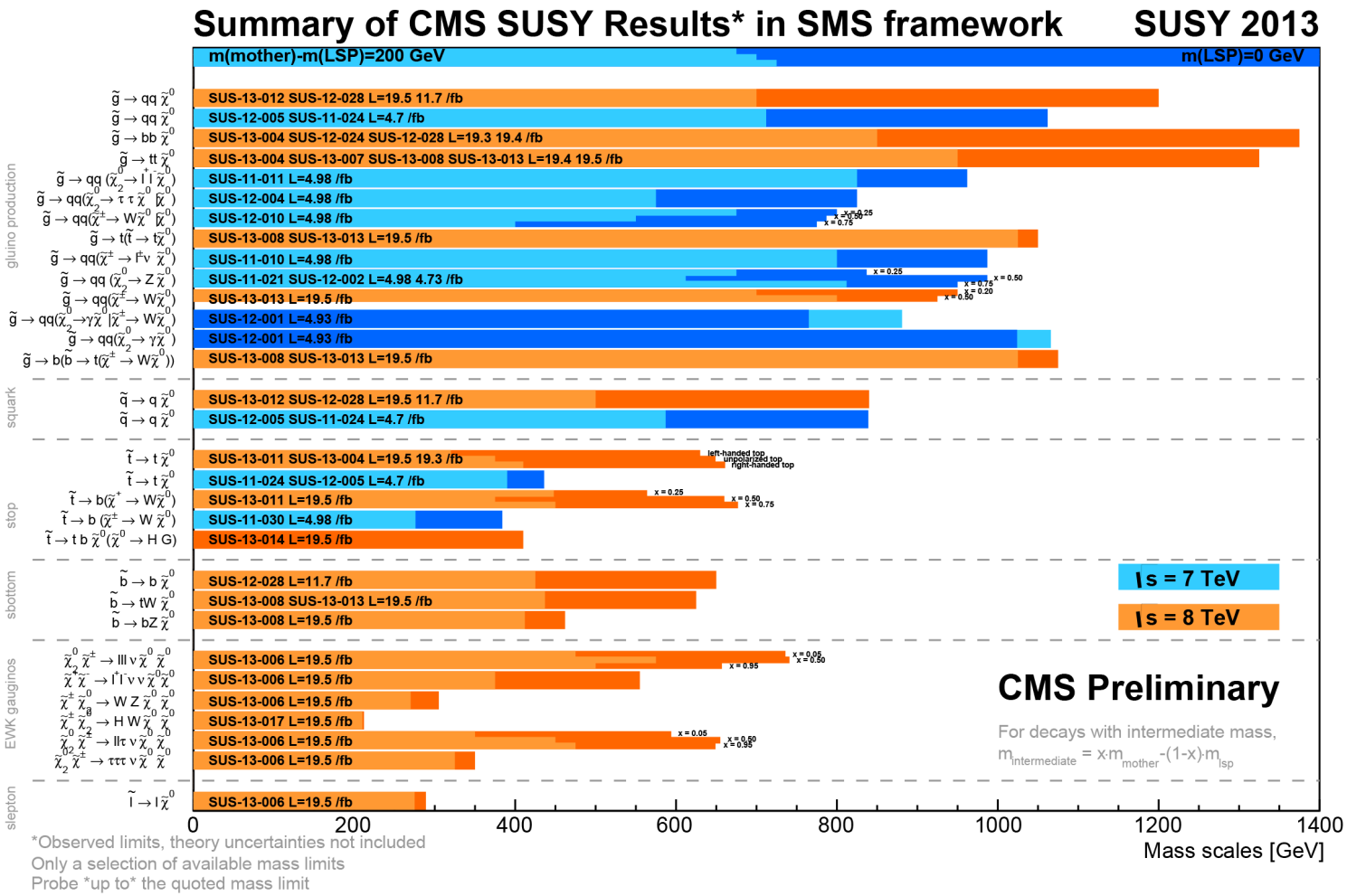


Figure 7-32. Limits set in selected CMS SUSY searches are shown, where some are based on 2011 data, and others on 2012 data. The limits represented by this thesis analysis are labeled SUS-13-013.

CHAPTER 8
CONCLUSIONS

The results of a search for a new physics signal involving events with 2 same-sign leptons, b-jets, and $\not\!\!\mathrm{E}_{\mathrm{T}}$ have been presented. This search provides the latest update to this topology to be performed on LHC data with the CMS experiment. It is carried out on the full 2012 LHC dataset comprised of 19.5 fb^{-1} with a collision energy of $\sqrt{\mathrm{s}} = 8$ TeV. The data are analyzed in exclusive signal regions formed by dividing into bins the discriminating variables $\mathrm{H_T}$, $\not\!\!\mathrm{E}_{\mathrm{T}}$, $\mathrm{N_{jets}}$, and $\mathrm{N_{b-jets}}$. The latter can assume values of 0, 1, and 2 or more, so signatures both with and without third-generation squarks can be probed. The backgrounds are predicted with well-established data-driven methods where possible and with Monte Carlo simulated events otherwise.

The analysis presented in Chapter 7 was one part in a three-part effort. All three analyses were combined to produce the limits shown in this thesis. No significant deviation from standard model expectation is observed for either the the high-$\mathrm{p_T}$ or low-$\mathrm{p_T}$ analyses. Therefore, some of the SUSY phase space is excluded at 95% C.L. in the context of simplified SUSY models, where $\tilde{\mathrm{g}}$ mass is excluded up to 1050 GeV and $\tilde{\mathrm{b}}$ mass is excluded up to 550 GeV.

In addition, a parametrization of the analysis acceptance is provided in Chapter 7 for others outside the CMS collaboration to use to check whether various models have been excluded by this search.

As one of the cleanest and lowest background channels for new physics searches at the CMS detector, the same-sign dilepton channel will continue to be watched closely. For as the collision energy of the LHC increases together with the instantaneous luminosity, the potential to discover SUSY will improve, the same-sign dilepton channel is still expected to be one of the first searches to show unambiguous evidence of SUSY if it is there.

APPENDIX A
VERTEXING TWO LEPTONS

```
// Create  SUSY Vertex from two muon tracks
// First create muons
TransientTrack tt1 = theTTBuilder->build( mu1->innerTrack() );
TransientTrack tt2 = theTTBuilder->build( mu2->innerTrack() );
ParticleMass m_mass = 0.1056583715;
float m_sigma = 0.0000000035;
float m_chi = 0.;
float m_ndf = 0.;

// Declare a particle factory
KinematicParticleFactoryFromTransientTrack pFactory;

// Put the muons in a vector
vector<RefCountedKinematicParticle> LepsToVertex;
LepsToVertex.push_back( pFactory.particle( tt1, m_mass, m_chi, m_ndf, m_sigma) );
LepsToVertex.push_back( pFactory.particle( tt2, m_mass, m_chi, m_ndf, m_sigma) );

// Declare the vertex fitter
KinematicParticleVertexFitter fitter;

// Do the fit
RefCountedKinematicTree LepsVertexFitTree;
LepsVertexFitTree = fitter.fit( LepsToVertex );
RefCountedKinematicParticle SusyCand = LepsVertexFitTree->currentParticle();
RefCountedKinematicVertex SusyDecayVertex = LepsVertexFitTree->currentDecayVertex();

// Vertex getters
bool isValidVtx = SusyDecayVertex->vertexIsValid();
Global3DPoint posVtx = SusyDecayVertex->position(); //x,y,z position
GlobalError errVtx = SusyDecayVertex->error(); //error matrix
float chi2Vtx = SusyDecayVertex->chiSquared(); // SV chi^2
float ndfVtx = SusyDecayVertex->degreesOfFreedom(); //ndf=1 for all events

//code adapted from:
//  https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideKinematicVertexFit
//source code of fit() function at:
//  http://cmslxr.fnal.gov/lxr/source/RecoVertex/KinematicFit/src/
//  KinematicParticleVertexFitter.cc
//Vertex getters at:
//  http://cmslxr.fnal.gov/lxr/source/RecoVertex/KinematicFitPrimitives/
//  interface/KinematicVertex.h
```

APPENDIX B
COMPATIBILITY χ^2 FUNCTION

```
// Calculate chi2 of primary vertex and SUSY vertex, i.e. compatibility
float tools::chi2( const Global3DPoint p1, const GlobalError E1,
                   const Global3DPoint p2, const GlobalError E2 )
{
  AlgebraicSymMatrix33 E1m = E1.matrix();
  AlgebraicSymMatrix33 E2m = E2.matrix();

  // <dx> = <p1> - <p2>
  float dx[3];
  dx[0] = p1.x() - p2.x();
  dx[1] = p1.y() - p2.y();
  dx[2] = p1.z() - p2.z();

  // VD(inv) = (E1m + E2m)(inv)
  AlgebraicSymMatrix33 VD = E1m + E2m;
  VD.Invert();

  // chi^2 = <dx>(T) * VD(inv) * <dx>
  float chisquared = 0.;
  for( int i=0; i<3; i++ ) {
    for( int j=0; j<3; j++ ) {
      chisquared += VD(i,j)*dx[i]*dx[j];
    }
  }
  return chisquared;
}

// The function is called as such, where posVtx and errVtx are the SUSY
// vertex, obtained with vertex getters
float chi2_3D = chi2( PV, PVerr, posVtx, errVtx );
```

REFERENCES

[1] S. P. Martin, *A Supersymmetry primer*, hep-ph/9709356.

[2] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSYSMSSummaryPlots8TeV.

[3] Skhirtladze, Nikoloz, *Searching for Supersymmetry with Tri-Leptons and Missing Transverse Energy Using the CMS Experiment at the Large Hadron Collider*, *PhD thesis, University of Florida* (Dec, 2012).

[4] S. Dimopoulos and G. Giudice, *Naturalness constraints in supersymmetric theories with nonuniversal soft terms*, *Phys.Lett.* **B357** (1995) 573–578, [hep-ph/9507282].

[5] A. G. Cohen, D. Kaplan, and A. Nelson, *The More minimal supersymmetric standard model*, *Phys.Lett.* **B388** (1996) 588–598, [hep-ph/9607394].

[6] J. Bagger, J. L. Feng, and N. Polonsky, *Naturally heavy scalars in supersymmetric grand unified theories*, *Nucl.Phys.* **B563** (1999) 3–20, [hep-ph/9905292].

[7] D. Alves, N. Arkani-Hamed, S. Arora, et al., *Simplified models for lhc new physics searches*, *Journal of Physics G: Nuclear and Particle Physics* **39** (2012), no. 10 105005.

[8] L. Evans and P. Bryant, *LHC Machine*, *Journal of Instrumentation* **3** (2008), no. 08 S08001.

[9] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008), no. 08 S08004.

[10] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008), no. 08 S08003.

[11] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008), no. 08 S08002.

[12] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008), no. 08 S08005.

[13] LHCf Collaboration, *The LHCf detector at the CERN Large Hadron Collider*, *JINST* **3** (2008), no. 08 S08006.

[14] TOTEM Collaboration, *The TOTEM Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008), no. 08 S08007.

[15] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[16] The CMS Collaboration, *CMS Physics Technical Design Report, Volume II: Physics Performance*, *Journal of Physics G: Nuclear and Particle Physics* **34** (2007), no. 6 995.

[17] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideCategoryBasedElectronID.

[18] https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffs#22Jan2013_ReReco_of_2012_data_re.

[19] CMS Collaboration, *Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s}$=8 TeV*, *JHEP* **1401** (2014), no. arXiv:1311.6736 163.

[20] CMS Collaboration, *Search for new physics in events with same-sign dileptons and b jets in pp collisions at sqrt(s) = 8 tev*, arXiv:1212.6194.

[21] W. Adam, J. Andrea, C. Beluffi, et al., *Combination of b-tagging efficiency measurements in 2012 data at 8 tev pp collision*, *CMS Analysis Note* **AN-2012/470** (2012).

[22] **CMS** Collaboration, S. Chatrchyan et al., *Identification of b-quark jets with the CMS experiment*, *JINST* **8** (2013) P04013, [arXiv:1211.4462].

[23] D. Dobur et al., *Data-driven estimation of ttbar-like standard model background processes for same sign di-lepton signature*, *CMS Analysis Note* **AN-2010/378** (2010).

[24] D. Dobur et al., *Search for new physics with same-sign di-leptons, jets, and missing transverse energy using 2011 data*, *CMS Analysis Note* **AN-2011/184** (2011).

[25] R. Remington, D. Dobur, et al., *Search for new physics with same-sign di-leptons, jets and missing transverse energy with 4.7 fb^{-1} of cms data*, *CMS Analysis Note* **AN-2011/437** (2011).

[26] D. Dobur, R. Remington, et al., *Search for stop and sbottom production with same-sign di-leptons, b-tagged jets and met*, *CMS Analysis Note* **AN-2012/001** (2012).

[27] **CMS** Collaboration, *Search for new physics with same-sign isolated dilepton events with jets and missing energy*, *CMS Physics Analysis Summary* **CMS-PAS-SUS-11-010** (2011).

[28] CMS Collaboration, *Absolute luminosity normalization*, *CMS Detector Performance Summary* **CMS-DP-2011-003** (2011). The value of 4% is inflated to 6%, based on observations in 2011 data sample.

[29] C. Campagnari, A. George, F. Golf, et al., *Search for new physics in a final state with a same-sign lepton pair, jets, and missing transverse energy using the full 2012 cms dataset of cms*, *CMS Analysis Note* **AN-2013/120** (2013).

[30] CMS Collaboration, *Determination of jet energy calibration and transverse momentum resolution in cms*, *Journal of Instrumentation* **6** (2011), no. 11 P11002.

[31] M. Chen., "Lands: A statistical tool for calculating limits and significance." https://mschen.web.cern.ch/mschen/LandS/index.html, April, 2011.

[32] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SUSYSMSSummaryPlots8TeV.

BIOGRAPHICAL SKETCH

Lana Muniz grew up in Mt. Zion, IL with her parents and three siblings. From an early age, she was encouraged to pursue her chosen career in science. At 15, she moved to Aurora, IL to attend the Illinois Math and Science Academy, a public boarding school for bright and motivated high school students, founded by Leon Lederman, Nobel laureate in Physics. These formative years away from home afforded daily contact with the brightest minds and most interesting people, students and teachers, of her life. Lana went on to attend the University of Illinois at Urbana-Champaign, where she received a B.S. in Specialized Physics. She next worked at Fermilab as a student researcher, gaining experience in data analysis and an inside look at the exciting world of particle physics research. With the LHC poised to start operations, she was inspired to attend graduate school to study particle physics. She is very proud to be a part of this historic era in physics.